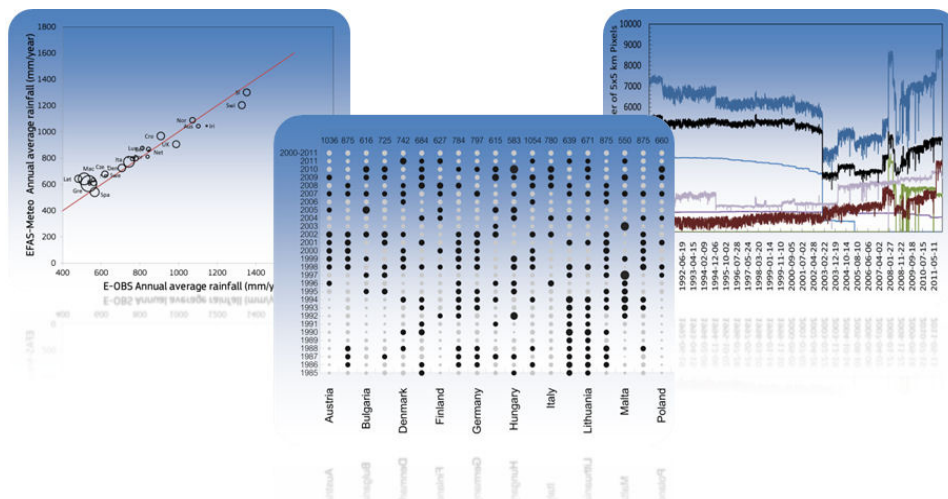
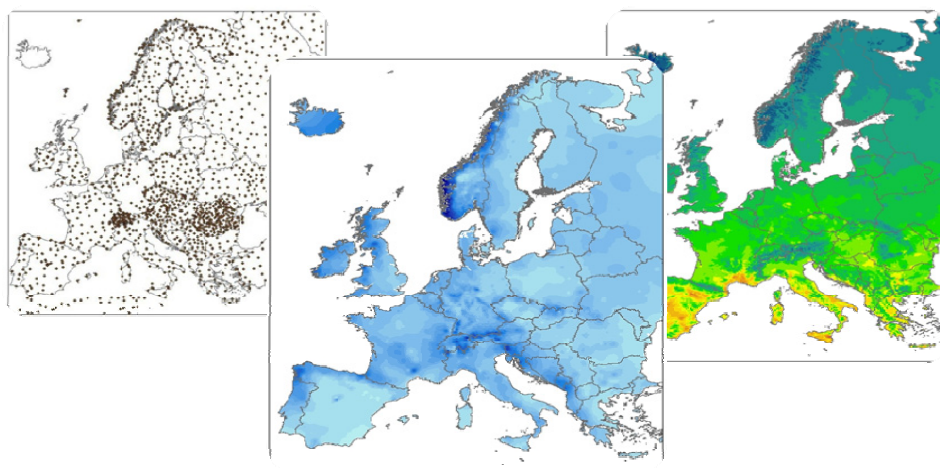


European  
Commission



# J R C T E C H N I C A L R E P O R T S

## EFAS-Meteo: A European daily high-resolution gridded meteorological data set for 1990 - 2011



V. Ntegeka, P. Salamon, G. Gomes, H. Sint, V. Lorini, M. Zambrano-Bigiarini, J. Thielen

**2013**

Report EUR 26408 EN

**European Commission**

Joint Research Centre

*Institute for Environment and Sustainability*

**Contact information**

Peter Salamon

Address: Joint Research Centre, Via Enrico Fermi 2749, TP 261, 21027 Ispra (VA), Italy

E-mail: [peter.salamon@jrc.ec.europa.eu](mailto:peter.salamon@jrc.ec.europa.eu)

Tel.: +39 0332 78 6013

<http://ies.jrc.ec.europa.eu/>

<http://www.jrc.ec.europa.eu/>

This publication is a Reference Report by the Joint Research Centre of the European Commission.

**Legal Notice**

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Europe Direct is a service to help you find answers to your questions about the European Union  
Freephone number (\*): 00 800 6 7 8 9 10 11

(\*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.  
It can be accessed through the Europa server <http://europa.eu/>.

JRC86388

EUR 26408 EN

ISBN 978-92-79-35006-1 (pdf)

ISSN 1831-9424 (online)

doi: 10.2788/51262

Luxembourg: Publications Office of the European Union, 2013

© European Union, 2013

Reproduction is authorised provided the source is acknowledged.

1	Introduction	4
1.1	Objective	4
2	Data Sources and Collection	5
2.1	EU-FLOOD-GIS database	5
2.2	JRC MARS database	5
2.3	Data extraction and interpolation	6
3	Meteorological variables	9
3.1	Precipitation	9
3.2	Minimum, maximum and average temperature	16
3.3	Calculated Radiation	19
3.4	Wind speed	21
3.5	Vapour pressure	21
3.6	Evapotranspiration	23
4	Analysis of trends	28
4.1	Trend analysis procedure	28
4.1.1	Annual temperature trends	28
4.1.2	Seasonal temperature trends	28
4.1.3	Annual precipitation trends	30
4.1.4	Seasonal precipitation trend analysis in extremes	31
5	Summary and conclusions	34
	Acknowledgements	34
	References	34
	Technical Appendix	37

# 1 Introduction

Data sets of spatially irregular meteorological observations interpolated to a regular grid are not only important for climate analyses but are also essential in order to derive climatologies for rainfall-runoff models which require meteorological data sets as input forcing. For example, in the European Flood Awareness System (EFAS (Thielen et al., 2009; Bartholmes et al., 2009) long term observed meteorological data are used to drive the hydrological model LISFLOOD (van der Knijf, 2010) to obtain long term time series of simulated discharges at a pan-European scale. Those long term time series of simulated “proxy” discharges can then be used for statistical analysis, e.g., to derive return periods or other time series derivatives.

There are several similar daily gridded data sets already available for Europe few of them contain the range of variables presented here or a similarly high spatial resolution. HadGHCND (Caesar et al., 2006) is a global gridded daily data set based upon near-surface maximum and minimum temperature observations. It spans the years 1946–2000 on a  $2.5^\circ$  latitude by  $3.75^\circ$  longitude grid. Furthermore, the JRC MARS-STAT database contains European meteorological observations interpolated onto a 50 km grid from 1975 up to present. Finally, the E-OBS gridded dataset covers the entire Europe with a grid resolution of  $0.25^\circ$  and spans a time from 1950 until 2012 (Haylock et al., 2008).

In this report, we present a comprehensive pan European high-resolution gridded daily data set (EFAS-Meteo) of precipitation, surface temperature (mean, minimum and maximum), wind speed, vapour pressure, calculated radiation and evapotranspiration (potential evapotranspiration, bare soil and open water evapotranspiration). The data set was created as part of the development of EFAS and has been continuously updated throughout the last years. Given the different variables, it was necessary to formulate procedures that would transform the data into spatial data with similar characteristics such as temporal and spatial grid resolutions. The PCRaster format (Karszenberg, 1996) was chosen for its ability to support modelling of physical processes at grid-level dimensions. A  $5 \times 5 \text{ km}^2$  grid resolution was chosen based on the set up of the pan European hydrological model used in EFAS. However, in some cases, a  $1 \times 1 \text{ km}^2$  grid resolution was used in some intermediate steps for a better representation of some processes (e.g., correcting temperature for elevation) before upscaling the maps to the final  $5 \times 5 \text{ km}^2$  grid resolution. It is anticipated that the data set will become an important basis for further climate analysis such as the establishment of decadal trends in Europe.

## 1.1 Objective

The principal objectives for the development of this long term, high resolution meteorological dataset are to produce a dataset of sufficient qualitative and quantitative accuracy and spatial variability to drive a European hydrological model with a spatial grid spacing of  $5 \times 5 \text{ km}^2$  to obtain skilful discharge simulations for river basins larger than  $1000 \text{ km}^2$ . The datasets are used as inputs to the calibration exercise as well as for establishing long-term discharge “proxy” climatologies.



## 2 Data Sources and Collection

The source data used to generate the EFAS-Meteo data set is collected and stored in two databases which are described briefly in the two following chapters.

### 2.1 EU-FLOOD-GIS database

The EU-FLOOD-GIS database is a data integration system for hydrological and meteorological data of various types and characteristics, e.g. time series and spatial data, which has been designed and developed as specific support to EFAS. Its first priority is the collection of both historic and real time meteorological and discharge/water level data. Currently data from, circa 4000 meteorological stations and 750 hydrological stations are being collected. The second priority is information on major lakes, reservoirs and structures and the third priority is remote sensing data such as weather radar products and satellite information. When available, other data relevant in the context of floods, e.g. flood extent vectors, are also collected and stored. The core solution for data collection and initial quality checks implemented for EU-FLOOD-GIS is based on Microsoft BizTalk. It provides a standards-based integration platform that combines messaging, web services, data transformation, and intelligent routing in an event-driven Service Oriented Application widely used to integrate and manage automated business processes. BizTalk has been customized from its default configuration to provide EU-Flood-GIS specific data plausibility, conversion, and processing functions before storing the output data in the EU-Flood-GIS data warehouse. EU-Flood-GIS BizTalk applications receive data files from provider organizations via FTP or Email. These raw data contain meteorological and/or hydrological data in several formats: SYNOP, XML and flat text. Furthermore, the meteorological and hydrological data are loaded using their original temporal resolution and type of measurement (e.g., instantaneous, average, accumulated, etc.) as transmitted by the different data providers.

Every single sample goes through several steps of processing (various validity checks, flagging, and insertion into RDBMS). The system is set-up such that the data transfer, processing and storage is automated and requires little intervention while at the same time ensuring that only data respecting the same format and standards are being loaded into the database. However, data control can obviously not be restricted to semantic standards but in order to ensure that EFAS starts up with the best initial conditions, the incoming data collection needs to undergo severe checks on content. Because of the importance of this issue, the collected data undergoes quality checks at several stages of the data collection and extraction process. They are implemented in different parts of the system workflow, based on the computational demands they pose to the system. While processing the data (MS2), the data records are checked against erroneous metadata such as not corresponding station identifiers, wrong date and time stamps, etc. These are filtered out before the loading procedure is started. Those having passed this initial check are being parsed into individual records and then checked against critical values and flagged accordingly as “green – acceptable”, “amber-dubious” and “red – unacceptable”.

Duplicates constitute a heavy overhead on the operations of the EU-FLOOD-GIS. In particular when collecting meteorological data, it often occurs that the same data are distributed through various providers, e.g. through the WMO GTS synoptic network, ECOMET networks and the National provider. Although this represents overlaps, the data collection becomes less dependent on local data transfer networks. In addition, duplicates also occur within the transmission of files. The WMO GTS synoptic data files regularly contain the same data several times to avoid loss of data. In the EU-FLOOD-GIS as many as 190000 duplicates out of 500000 records are at times recorded daily. Duplicates are stored in the database with a ranking flag, giving the local provider the highest rank.

One important feature of the Biztalk solution is its “Business Activity Monitor” (BAM) implementation. BAM implementation in EU-FLOOD-GIS allows monitoring parameters like number of files received, number of samples received, processed, processing status and errors (e.g.), data flags and aggregations of all these variables per time interval/provider/data type. Further detailed information on the EU-FLOOD-GIS database can be found in Thielen et al. (2012).

### 2.2 JRC MARS database

The JRC MARS database contains, amongst other data, daily meteorological data, which has been specifically designed and developed in support to agricultural applications, specifically the Crop Growth Monitoring System (CGMS) which is developed and managed by the JRC. The meteorological station data in the MARS database are available in near real time, i.e., maximum 1 day after the acquisition time and the data consists of station information, the raw daily meteorological data and possibly of the processed meteorological daily data. In the JRC MARS database some of the meteorological data are purchased directly from various national meteorological services, others are acquired via the Global Telecommunication System (GTS). As the data are obtained from a variety of different sources, considerable pre-processing is necessary to convert them to a standard format. Two different procedures are applied for distinct subsets

of the data set. The historic data are ordered directly from national meteorological services and are then converted into consistent units and checked on realistic values. The near real time data are pre-processed and quality checked using the AMDAC software package (MeteoConsult, 1991) which extracts, decodes and processes the GTS data. After decoding, the following data are checked for consistency and errors: air temperature, dew-point temperature (humidity), pressure at sea level, wind speed, amounts of precipitation, clouds, and sunshine duration. This error checking compares each observation with the corresponding values of the surrounding stations and compares that particular observation with observations at other times in the same day at the same station. Obvious errors in the observations are corrected automatically and a message is written to a log file; other errors are flagged for possible correction by an operator (Burrill and Vossen, 1992). Finally, the data are converted into daily values. This comprises the selection of minimum and maximum temperature, the aggregation of the rainfall, cloud cover and sunshine duration, the calculation of mean vapour pressure etc.

The processed daily meteorological data consists of 30 meteorological parameters including various cloud cover indicators, air temperature, vapour pressure, wind speed and rainfall. Because European stations follow different measurement schemes many records contain blank fields for parameters which are never registered. Stations often also include blank fields for parameters which were not available for limited periods. However, the stations selected for inclusion in the database are those which normally report at least the minimum and maximum daily air temperature, rainfall, wind speed, vapour pressure (or humidity) as well as either global radiation, sunshine hours or cloud cover (Burrill and Vossen, 1992).

### 2.3 Data extraction and interpolation

Data from the JRC MARS and the EU-FLOOD-GIS database is extracted using the 2Map application (operation described in Table 1), which has been developed in house. 2Map is written in Python following OOAD principles. 2Map directly queries both databases, EU-FLOOD-GIS and JRC MARS, and it offers various options of execution. When executed, the program reads daily data - for the desired variable and date - from the operational table, which is being preloaded with possibly aggregated data via scheduled PL/Sql procedures. For example, daily precipitation data can comprise either directly daily accumulated data values or is constructed by aggregating 12 hourly, 6 hourly, 3 hourly, hourly, 15 min data to daily accumulates. However, this aggregation is currently only performed for the variable precipitation. For all other variables currently only data is used which has a daily time stamp in the EU-FLOOD-GIS database. The output data includes a raster map (PCRaster format).

**Table 1: Input arguments that are available when executing the 2Map application**

Parameter	Possible values																						
<p>-f, --from:</p> <p>Start date of extraction, issued with the format YYYYMMDD, which will be a date at 0 AM. Internally, the from_date will be set to from_date = from + start_hour_for_aggragation. The value start_hour_of_aggregation depends on the variable. This setting can be modified directly in the XML configuration file.</p> <p>-n, --steps:</p> <p>Number of steps</p> <p>-s, --steplength:</p> <p>Step length in hours</p> <p>-v, --vars:</p> <p>A list of variable codes, comma separated</p>	<p>A suitable date in the format YYYYMMDD. e.g.: 20110321</p> <p>A non negative integer greater than 0. e.g. 1</p> <p>24</p> <table border="1"> <thead> <tr> <th>Variable Code</th><th>Variable Name</th></tr> </thead> <tbody> <tr> <td>ta</td><td>Temperature daily average</td></tr> <tr> <td>tn</td><td>Temperature daily minimum</td></tr> <tr> <td>tx</td><td>Temperature daily maximum</td></tr> <tr> <td>td</td><td>Dew Point temperature</td></tr> <tr> <td>pr</td><td>Precipitation daily</td></tr> <tr> <td>qa</td><td>Discharge daily average</td></tr> <tr> <td>cc</td><td>Cloud Coverage, daily</td></tr> <tr> <td>ws</td><td>Wind Speed, daily average</td></tr> <tr> <td>pd</td><td>Vapour pressure</td></tr> <tr> <td>rg</td><td>Calculated radiation</td></tr> </tbody> </table>	Variable Code	Variable Name	ta	Temperature daily average	tn	Temperature daily minimum	tx	Temperature daily maximum	td	Dew Point temperature	pr	Precipitation daily	qa	Discharge daily average	cc	Cloud Coverage, daily	ws	Wind Speed, daily average	pd	Vapour pressure	rg	Calculated radiation
Variable Code	Variable Name																						
ta	Temperature daily average																						
tn	Temperature daily minimum																						
tx	Temperature daily maximum																						
td	Dew Point temperature																						
pr	Precipitation daily																						
qa	Discharge daily average																						
cc	Cloud Coverage, daily																						
ws	Wind Speed, daily average																						
pd	Vapour pressure																						
rg	Calculated radiation																						
<p>-o, --output:</p> <p>Output folder for maps/files</p> <p>-c, --configuration:</p> <p>The desired configuration to use between those which are present in the configuration file.</p> <p>-m, --outputmode:</p> <p>The desired output format.</p> <p>-e, --efgOnly</p> <p>When this option is set, the program will produce maps out of EFG dataset only in addition to those ones created with the joint of EFG and MARS datasets.</p> <p>-z, --stationId</p> <p>The station id for which you want to extract data. When this option is used, the output mode will be set automatically to 'tabular'.</p> <p><b>-h, --help:</b></p> <p><b>show the help and exit</b></p>	<p>A well formed pathname. If it doesn't exist it'll be created</p> <p>An existing configuration in application-configuration.xml file</p> <p>maps, files, tabular</p> <p>False, True</p> <p>An existing station_id.</p>																						

The option -outputmode provides currently three different formats:

**files:** For each variable and for each step, a text data file will be saved in the output folder in the X, Y, Z, SITE\_NAME, VALUE, AGGREGATION\_TYPE\_CODE format.

**tabular:** The program will produce a single file in tabular format, that means: each row per day and each column per station.

**maps:** The program will produce PCRaster maps out of daily data. For each variable and for each step, in the output folder you will find interpolated maps, maps of points (without interpolation, in order to better check the coverage), and a text data file in the X, Y, VALUE format. This option was used to produce the current dataset and is used also in the operational EFAS system to calculate the initial conditions for the water balance.

All meteorological variables are interpolated using an inverse distance scheme with a weight of  $d^{-2}$  and a maximum number of stations used in the computation of an interpolated pixel of 5. Throughout the development of the dataset various kriging approaches instead of inverse distance interpolation have been tested using the gstat R package (as wrapped by the hydroTSM R package), which is commonly used for such purposes. The following approaches were evaluated: kriging with external drift (KED) for temperature and precipitation, and a combined indicator kriging and ordinary kriging for precipitation. Concerning kriging with external drift for temperature, no correlation between height and temperature could be detected at the continental scale (temperature measurements at the same altitudes in northern Scandinavia obviously have no correlation to temperature measurements at similar altitudes in the Mediterranean). Furthermore, correlation between daily precipitation and elevation was found to be low or non-existent at continental scale when all the available gauging stations are used for computing the correlation. Hence, KED was discarded as a plausible approach. For precipitation, a second approach was investigated which has been applied by various authors for large scale interpolation (e.g., Hofstra et al., 2008). In this approach indicator kriging, with daily variogram fitting, is first applied to generate a Boolean map showing pixels with rain and no rain, by using a probability of rain equal to 0.4 to discriminate between them. Then, ordinary kriging is used to interpolate the amount of precipitation only onto the pixels which have been previously denoted as pixels with rain. This method showed promising results especially for rainy days. Nevertheless, when used to generate a long-term historic data set some of the precipitation maps had unrealistically high values of precipitation. This was caused principally by the automatic fitting of the variogram to the measured data, which was not stable enough to produce a realistic variogram for each day, especially when nearby stations registered very different amounts of precipitation, or when the “screening effect” was important. The aforementioned issue may be partially due to the fact that different data sources may have different spatial coordinates for the same gauging station, and –for some cases– even with different amounts of precipitation. This problem could not be solved, even by limiting the variogram fitting parameters (e.g., fixing the nugget, using a maximum searching radius of 250 km, and a maximum amount of neighbours equal to 7), or by using an especially robust variogram fitting (Cressie). Thus, keeping in mind that the interpolation algorithm chosen needs to be absolutely stable, because of its use in the operational EFAS system, it was decided to maintain the inverse distance method.

### 3 Meteorological variables

Given that a 30-year period is recommended by climatologists for characterizing the climate of an area, it was worthwhile to investigate the existing dataset beginning from the 1980s. The aim was to extract the longest reliable time series such that the climatology of the region could be well represented. The meteorological data were extracted initially for the period 1985–2011. However, due to data quality issues for precipitation, which will be outlined in the following sections, the final dataset was extracted for the period 01/01/1990 until 31/12/2011. The year 2012 was not included as it was not yet a complete year. Efforts will be made to update the data in the future. Spatial comparisons of annual rainfall over the region were made from year to year. Additionally, station density time series plots were made to help in explaining the observed spatial anomalies. By comparing spatial maps and time series plots, it was possible to discern whether the anomalies could be attributed to the spatial interpolation or the quality of the station data.

Table 2 provides a detailed definition of the different meteorological variables extracted for this dataset. For precipitation the definitions are not limited to the ones described in the table, however, due to the various time intervals over which measurements are recorded by different providers. For instance, intervals such as 7 UTC–7UTC, 6:30UTC–6:30UTC to mention but two were included in the data set. This was done because we considered that a higher station density is more important for capturing better the precipitation fields, than the exact agreement with WMO standards.

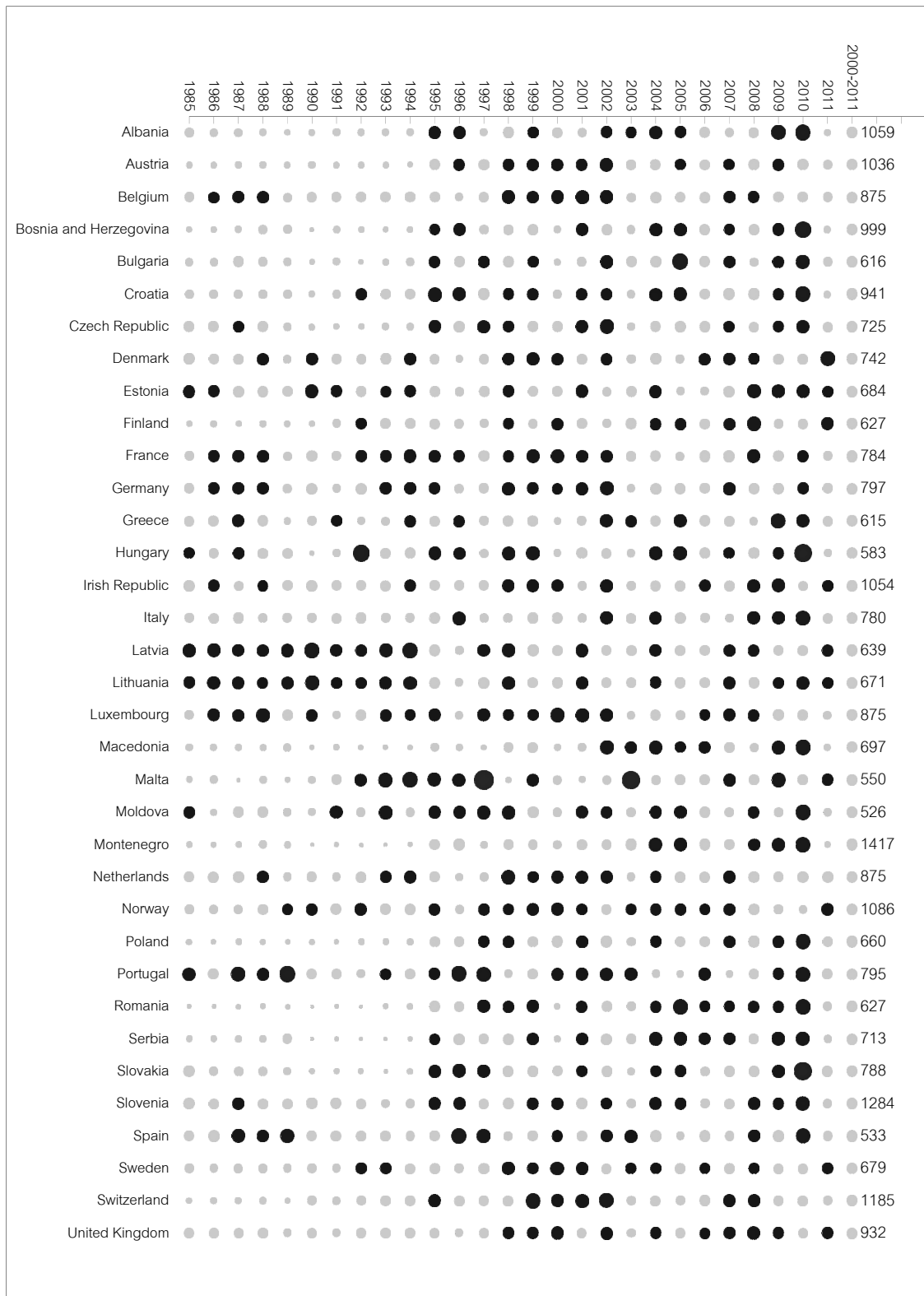
**Table 2: Definition of meteorological variables as used in the EFAS-Meteo dataset**

pr	Daily precipitation (mm) between 6 UTC on the day specified and 6 UTC on the next day
tn	Daily minimum temperature (°C) between 18 UTC and 6 UTC (i.e. during the preceding night) at 2m
tx	Daily maximum temperature (°C) between 6 UTC and 18 UTC (i.e. during daytime) at 2m
ta	Daily mean temperature (°C) is calculated using $ta = (tx + tn) / 2$
ws	Mean daily wind speed at 10 metres (m/s) calculated from 3-hourly observations (0–24 UTC)
pd	Mean daily vapour pressure (hPa)
cr	Calculated radiation (KJ/m <sup>2</sup> /day)
e0	Penman potential evaporation from a free water surface (mm/day)
et	Penman potential transpiration from a crop canopy (mm/day)
es	Penman potential evaporation from a moist bare soil surface (mm/day)

#### 3.1 Precipitation

Precipitation is arguably the most crucial hydrological variable for hydrological forecasting; other crucial factors are temperature and evapotranspiration. To check the precipitation data quality, two procedures were employed. The first involved checking for annual consistency within the EFAS-Meteo dataset. By comparing year to year anomalies, possible errors can be captured manifested as large deviations from an expected annual estimate. The second approach involved comparing two datasets. The idea was to compare the long term averages for two datasets to ascertain major differences. The E-OBS observational dataset (Haylock et al., 2008) was selected for comparison. Annual precipitation values for the E-OBS daily gridded dataset (version 6.0) at 25 x 25 km<sup>2</sup> resolution were compared to the EFAS-Meteo values (5 x 5 km<sup>2</sup> resolution).

Previous analyses (not shown here) and preliminary tests suggested that there might be some problems concerning a significant underestimation of precipitation, especially for the time periods before 1995 (Figure 1). Therefore, the period 2000–2011 was used as a reference period (most recent decade) for estimating the expected annual precipitation. Furthermore, as most of the national/regional data providers included in this dataset have started to provide data regularly from 2000 onwards a reasonably high station density throughout most of Europe is available. Figure 1 shows the relative fraction of annual precipitation volumes as compared to the period 2000–2011. It can be observed that the annual volumes for Poland, Macedonia, Romania, and Montenegro between 1985–1994 are unrealistically low (dry) compared to the period 2000–2011, suggesting a systematic underestimation of precipitation for those regions. Furthermore, Finland exhibits similar dry patterns for the years 1985–1990.



**Figure 1 Relative fraction of annual precipitation volumes as compared to the period 2000-2011. The black shade implies wetter than the base period, while the grey shade implies drier than the base period. The size of each dot represents the magnitude. The right numeric labels are average annual volumes for 2000-2011 and are given for guidance in estimating the annual volumes for other years**

In addition, further checks were made comparing the annual average precipitation derived for periods from 1990-2011 and 1995-2011. Table 3 shows the annual averages for the EFAS-Meteo and the E-OBS datasets compared with annual average precipitation as reported by Eurostat (database name: env\_watq1a from epp.eurostat.ec.europa.eu). Clearly, some countries illustrate a large change in annual average precipitation when using a reference period that starts only from 1995-2011 instead of 1990-2011 compared to Eurostat. It is evident that for some regions indeed a significant underestimation of total annual precipitation recorded in the time period before 1995 exists. As similar discrepancies are also observed in the E-OBS dataset this suggests that this issue is not related to the different spatial

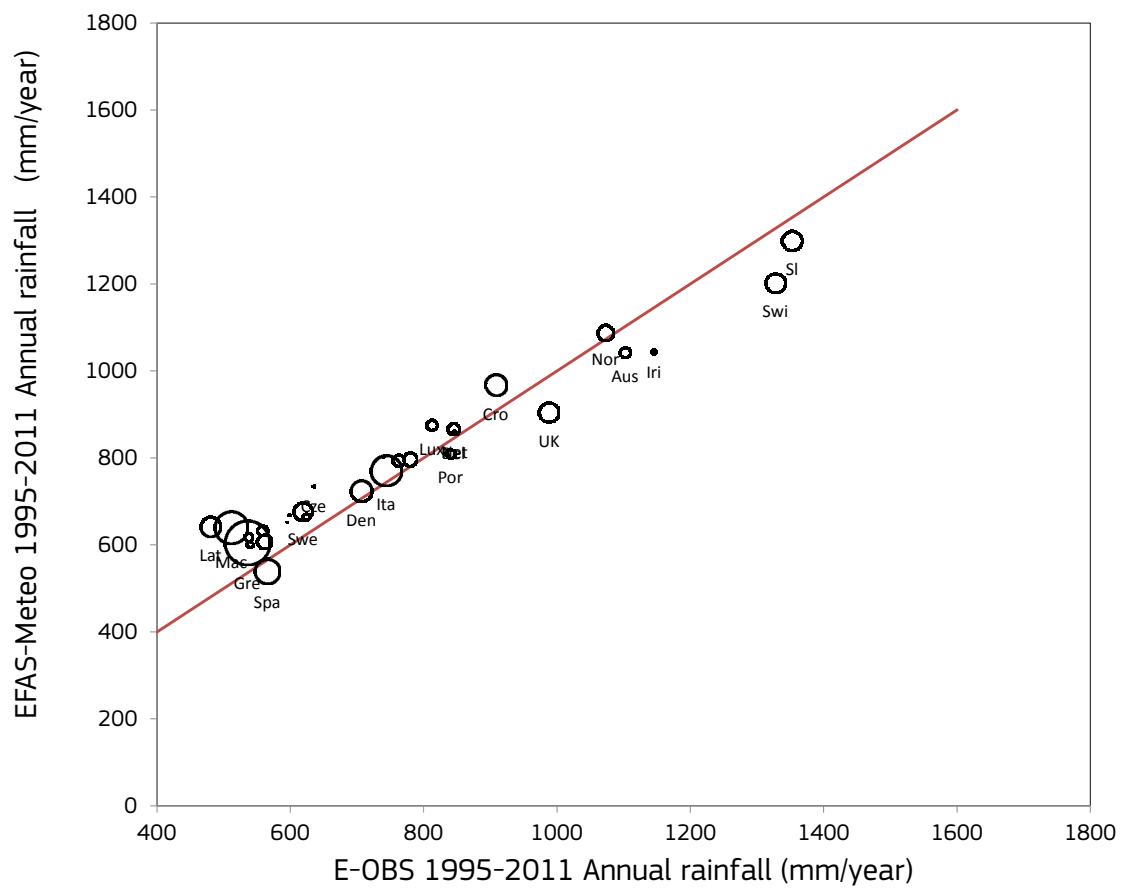
resolution or interpolation methods used in the two datasets. Instead, it appears that this systematic underestimation is caused by either a very low station density (see for example Bulgaria or Austria in Figure 4) or by a problem with the actual raw data of some of the providers.

Although being aware of those problems for some specific countries, we have chosen the data period 01/01/1990 until 31/12/2011 for our final dataset principally for two reasons: (1) Time series analysis of the EFAS climatology requires a certain minimum length of the meteorological dataset; (2) Data quality in the remaining regions seems to be good also for the period 1990 -1994. It is notable that the weighted average deviation of the EFAS-Meteo from Eurostat, which is calculated based on country size, is around +13% for 1990-2011 and +10% for 1995-2011.

**Table 3: Comparison between Eurostat and the two datasets, EFAS-Meteo and E-OBS. The magnitude differences are computed for two periods, 1990-2011 and 1995-2011**

	>=20years (end 2009)		1990-2011		1995-2011		1990-2011		1995-2011	
	Eurostat		E-OBS	EFAS-Meteo	E-OBS	EFAS-Meteo	E-OBS/Eurostat (%)	EFAS-Meteo/Eurostat(%)	E-OBS/Eurostat(%)	EFAS-Meteo/Eurostat(%)
Austria	1167.08		1086.21	940.03	1102.18	1042.49	6.93	19.45	5.56	10.68
Belgium	943.09		835.54	854.87	845.17	866.55	11.40	9.35	10.38	8.12
Bulgaria	618.07		508.12	560.84	537.47	617.98	17.79	9.26	13.04	0.02
Croatia	1116.86		896.74	928.07	908.31	967.64	19.71	16.90	18.67	13.36
Czech Republic	693.60		614.87	667.63	635.17	734.81	11.35	3.74	8.42	-5.94
Denmark	852.21		702.31	725.39	706.67	723.77	17.59	14.88	17.08	15.07
Estonia	641.87		616.48	681.06	598.27	668.31	3.95	-6.11	6.79	-4.12
Finland	659.55		561.70	590.76	561.35	607.46	14.84	10.43	14.89	7.90
France	884.23		787.94	797.55	779.82	796.91	10.89	9.80	11.81	9.88
Germany	858.61		761.84	792.19	762.74	793.90	11.27	7.74	11.17	7.54
Greece	869.04		520.88	589.11	535.40	604.91	40.06	32.21	38.39	30.39
Hungary	599.04		528.69	584.25	539.64	601.33	11.74	2.47	9.92	-0.38
Irish Republic	1137.56		1173.16	1035.74	1145.64	1044.18	-3.13	8.95	-0.71	8.21
Italy	980.86		741.52	751.00	743.79	771.08	24.40	23.43	24.17	21.39
Latvia	662.39		522.21	673.24	480.64	641.86	21.16	-1.64	27.44	3.10
Lithuania	679.82		629.31	685.22	623.68	664.10	7.43	-0.79	8.26	2.31
Luxembourg	781.58		804.43	866.32	812.29	875.60	-2.92	-10.84	-3.93	-12.03
Macedonia	771.33		489.31	573.05	511.54	640.72	36.56	25.71	33.68	16.93
Netherlands	830.66		840.28	854.03	846.28	858.87	-1.16	-2.81	-1.88	-3.40
Norway	1225.04		1079.29	1090.19	1072.94	1087.33	11.90	11.01	12.42	11.24
Poland	620.10		572.64	584.25	595.47	652.12	7.65	5.78	3.97	-5.16
Portugal	891.57		804.32	786.52	840.42	809.69	9.79	11.78	5.74	9.18
Romania	647.83		535.69	547.63	558.71	632.60	17.31	15.47	13.76	2.35
Slovakia	763.04		717.67	708.21	740.23	802.99	5.95	7.19	2.99	-5.24
Slovenia	1566.73		1356.59	1258.71	1352.61	1299.37	13.41	19.66	13.67	17.06
Spain	683.92		547.66	528.64	565.62	539.57	19.92	22.71	17.30	21.11
Sweden	752.93		617.04	664.61	618.86	676.46	18.05	11.73	17.81	10.16
Switzerland	1494.31		1328.04	1158.95	1327.88	1201.99	11.13	22.44	11.14	19.56
United Kingdom	1119.08		1003.28	882.62	988.01	904.71	10.35	21.13	11.71	19.16
Weighted average							14.71	13.10	13.92	10.15

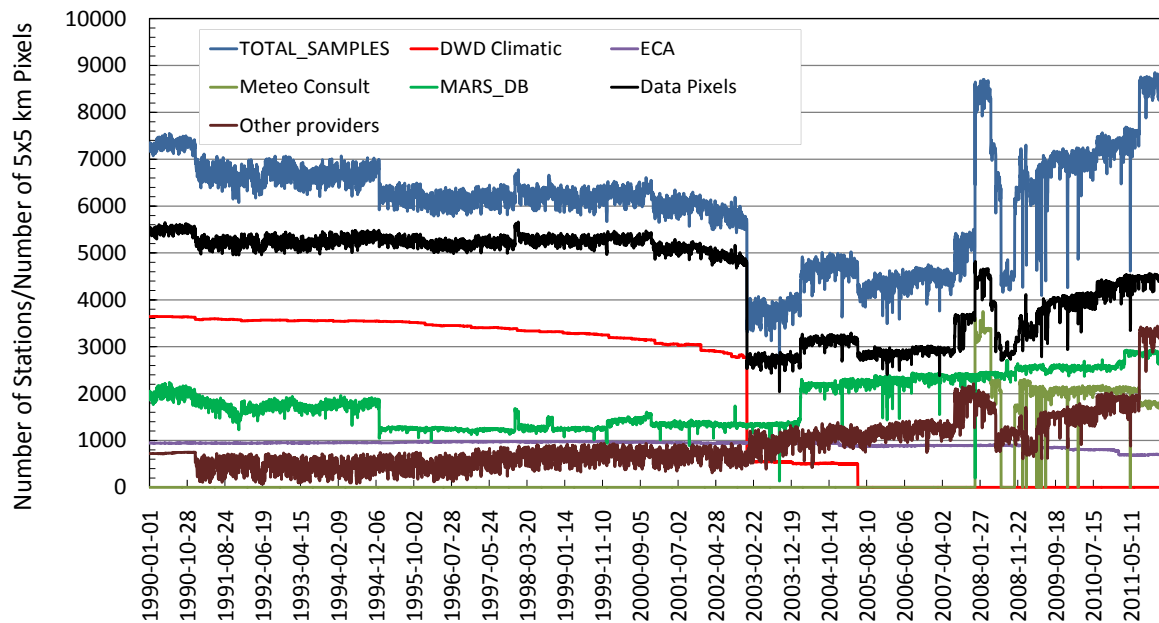
Figure 2 shows the annual average precipitation from the two datasets with the bubbles size representing the average percentage difference meaning that a smaller bubble size indicates better agreement with Eurostat. From this illustration it is found that especially countries such as Greece, Italy, Spain, United Kingdom, Macedonia, Croatia require further scrutiny to establish the source of the disparities. However, it may be possible that for some countries the discrepancies could also come from the Eurostat database especially where the two databases (E-OBS and EFAS-Meteo) tend to agree. For example, Greece and Italy appear to have similar large deviations from the Eurostat database. Whether this occurs due to inconsistencies in both EFAS-Meteo and E-OBS remains to be explored.



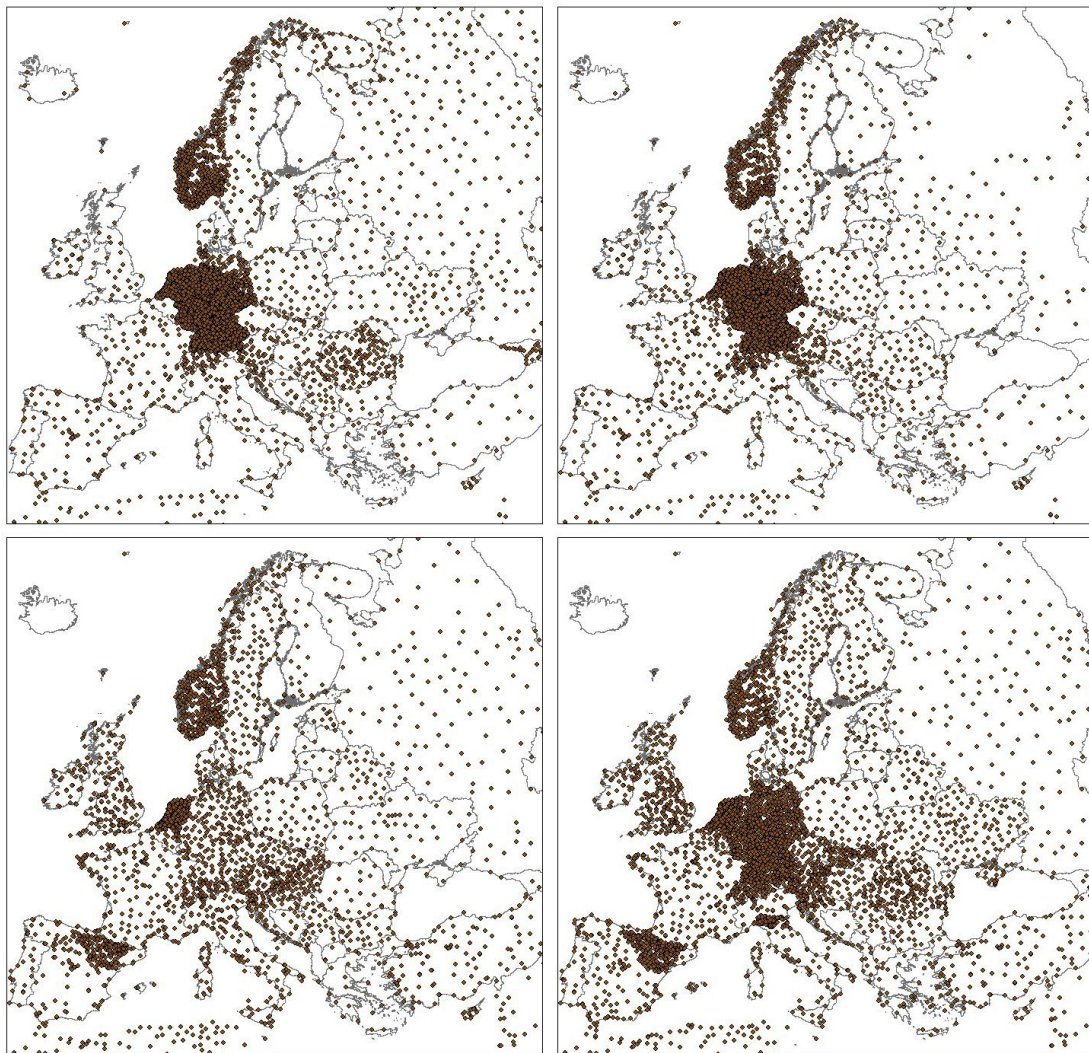
**Figure 2 Comparing annual average rainfall for E-OBS and EFAS-Meteo with Eurostat for different EU countries. The bubble size corresponds to the magnitude of divergence from the Eurostat annual volumes for the period 1995-2011. The bisector is included to illustrate the correlation**



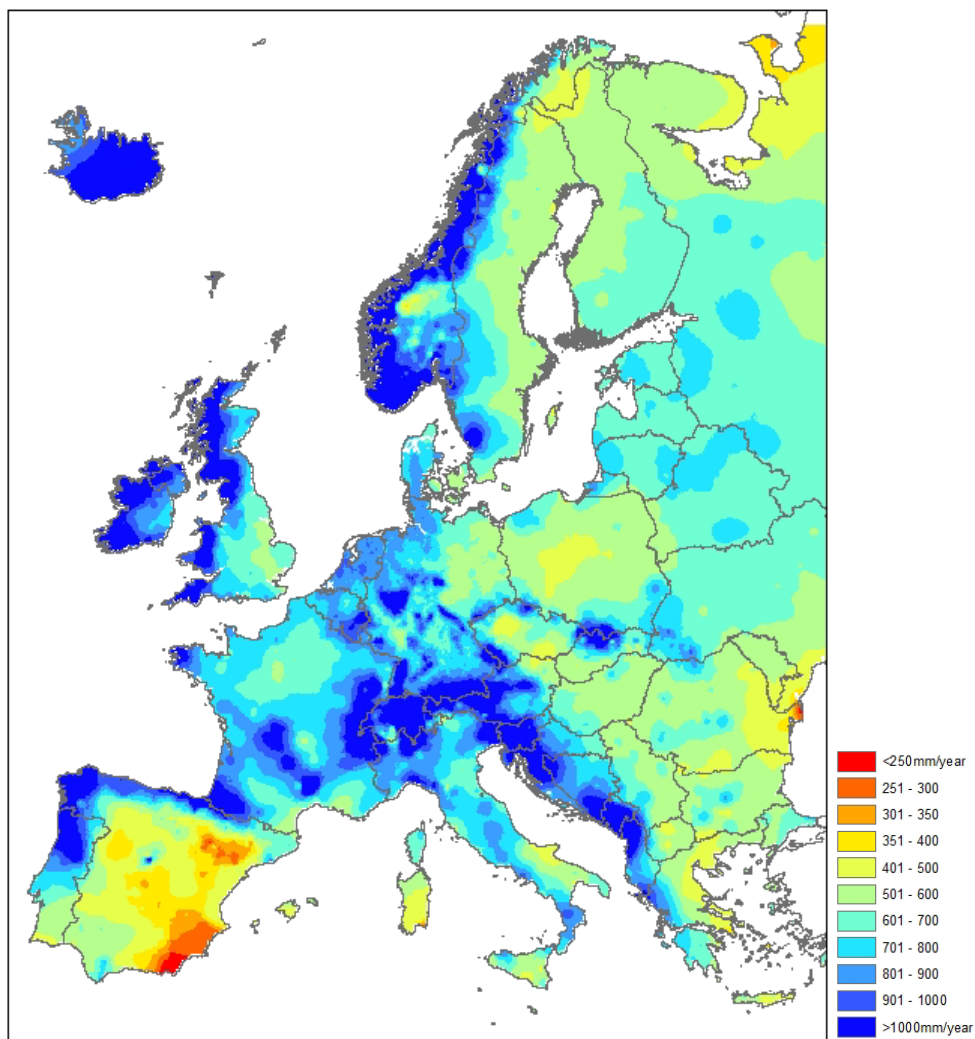
Figure 3 shows the temporal series of precipitation stations from 1990 to 2011. On average more than 6000 stations for daily precipitation are available. The main providers are DWD climatic, ECA and JRC MARS with each providing over 1000 stations. Between 2003 and 2008 there is a noticeable decline of available stations to around 4000 with a rise around 2007, which is principally caused by a drop in stations by the provider DWD Climatic. Note that this drop is not severe as to adversely affect the quality of the daily rainfall in Germany; the spatial distribution remains quite good for Germany despite the drop as shown in Figure 4 (31 Dec 2005). Figure 4 displays the spatial distribution of observation stations for selected days. It is evident from the station distribution that Germany as well as other national/regional providers such as Norway, Austria, Slovakia, Ebro river basin (Spain) etc. have a high station density (e.g., over 910 stations for Germany for 31/12/2011). The spatial annual average rainfall as derived from the time period 1990 – 2011 is displayed in Figure 5. Monthly average precipitation maps can be found in the appendix.



**Figure 3 Time series of available precipitation stations. The black time series shows the number of 5x5 km<sup>2</sup> pixels in the PCRaster map with one or more observations**



**Figure 4 Spatial distribution of available precipitation observations for 31 Dec 1990 (upper left), 31 Dec 1995 (upper right), 31 Dec 2005 (lower left), and 31 Dec 2011 (lower right)**

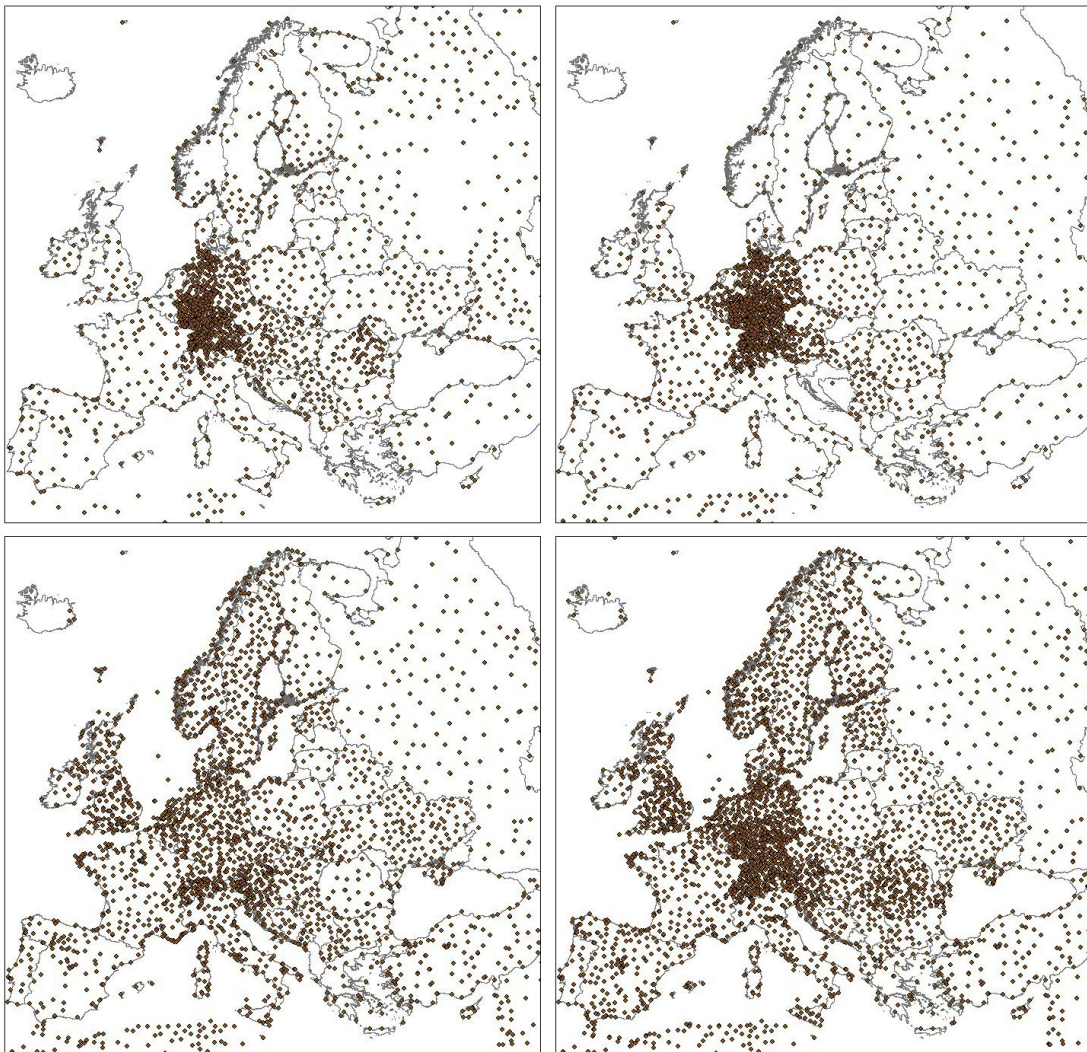


**Figure 5 Annual average precipitation for period 1990-2011 from the EFAS-Meteo dataset**

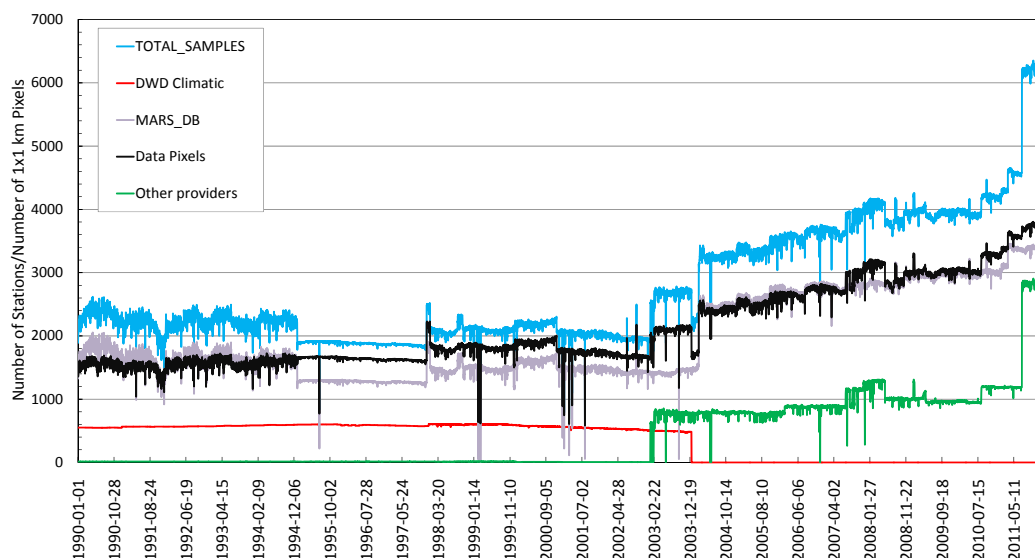


### 3.2 Minimum, maximum and average temperature

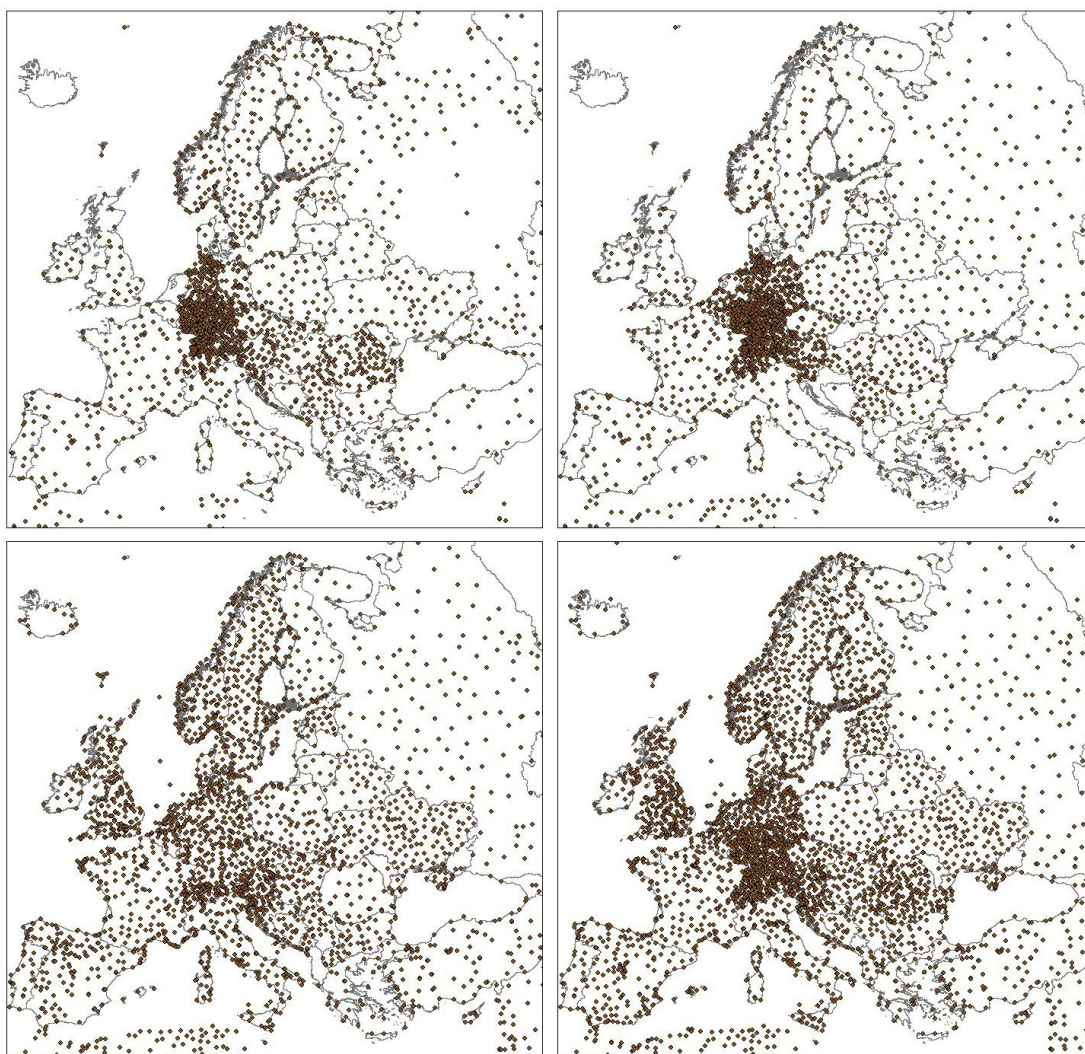
Unlike precipitation, temperature measurements were corrected for elevation using the following approach: First minimum and maximum temperature measurements were all brought to 0m height using a 1x1 km<sup>2</sup> resolution digital elevation map and a correction factor of -0.006 (°C/m). Then stations were resampled to 5x5 km<sup>2</sup> resolution, interpolated and then corrected using the 5x5 km<sup>2</sup> digital elevation map. Daily mean temperature was then calculated using the formula  $t_a = (t_x + t_n)/2$ . Monthly temperature maps can be found in the appendix. Figure 7 shows the number of 1x1 km<sup>2</sup> pixels having at least one or more minimum temperature observation. Figure 6 displays the spatial station distribution for selected days 31/12/1990, 31/12/1995, 31/12/2005 and 31/12/2011. The spatial distribution of the available stations for maximum temperature and the temporal time series are illustrated in Figures 8 and 9.



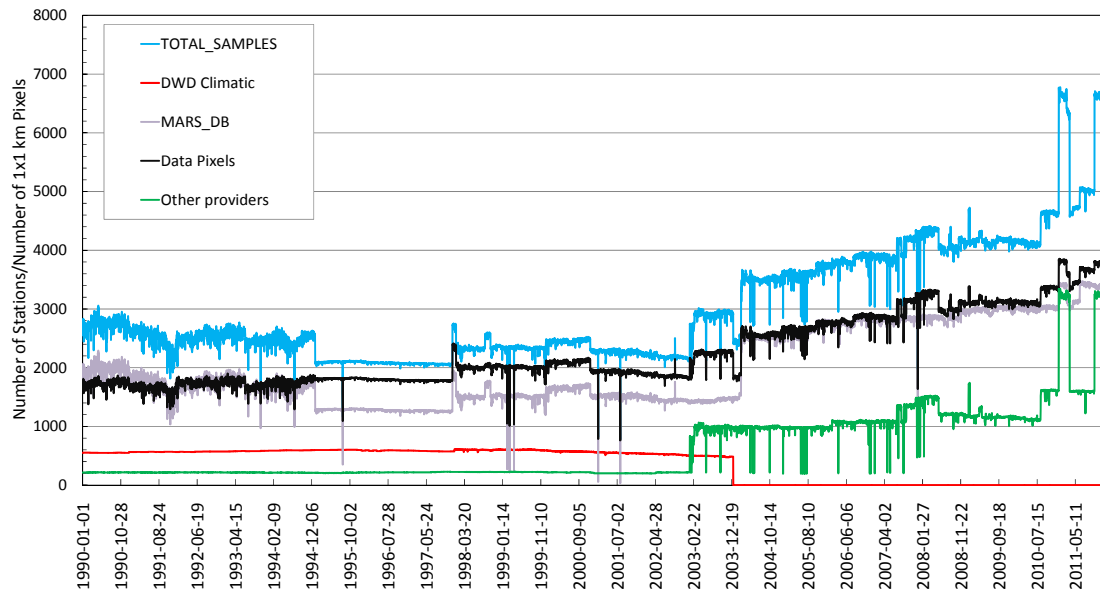
**Figure 6 Spatial distribution of available minimum temperature observations for 31 Dec 1990 (upper left), 31 Dec 1995 (upper right), 31 Dec 2005 (lower left), and 31 Dec 2011 (lower right)**



**Figure 7 Time series of available minimum temperature time series for the different data sources. The black time series shows the number of 1x1 km<sup>2</sup> pixels in the PCRaster map with one or more observations**

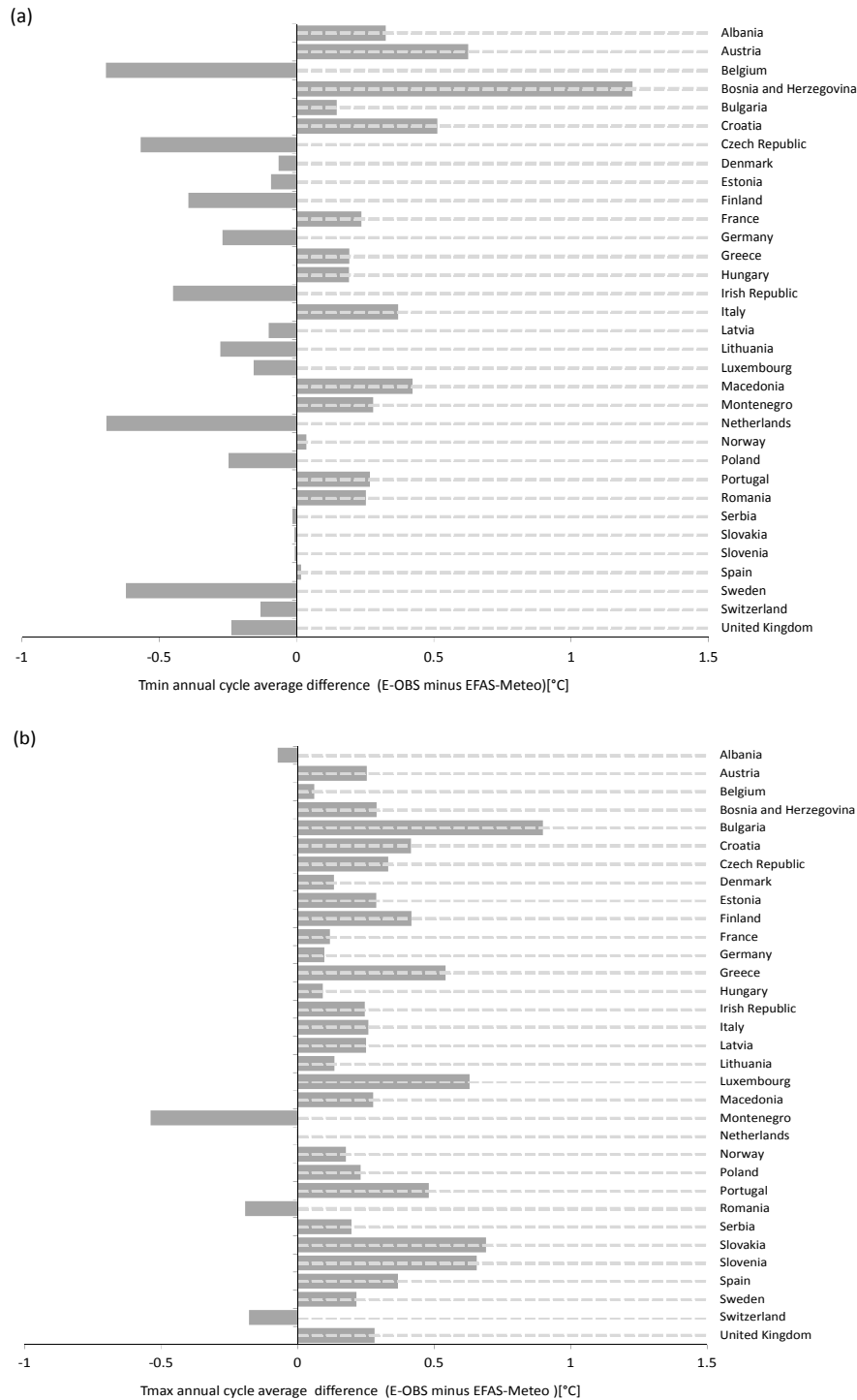


**Figure 8 Spatial distribution of available maximum temperature observations for 31 Dec 1990 (upper left), 31 Dec 1995 (upper right), 31 Dec 2005 (lower left), and 31 Dec 2011 (lower right)**



**Figure 9 Time series of available maximum temperature time series for the different data sources. The black time series shows the number of 1x1 km pixels in the PCRaster map with one or more observations**

A comparison with the E-OBS dataset revealed that the minimum and maximum temperatures were close to the EFAS-Meteo temperatures. The mean annual cycle temperature difference was calculated as the average deviation of the annual cycle (averaged over 1990-2011) of the E-OBS from the EFAS-Meteo dataset. It was found that for minimum temperature it was not evident which dataset was warmer or colder; the average deviation was close to 0°C. However, for maximum temperature, it was clear that E-OBS showed a warmer climate than EFAS-Meteo with an average of around 0.24°C. The annual temperature cycles per country can be found in the Appendix

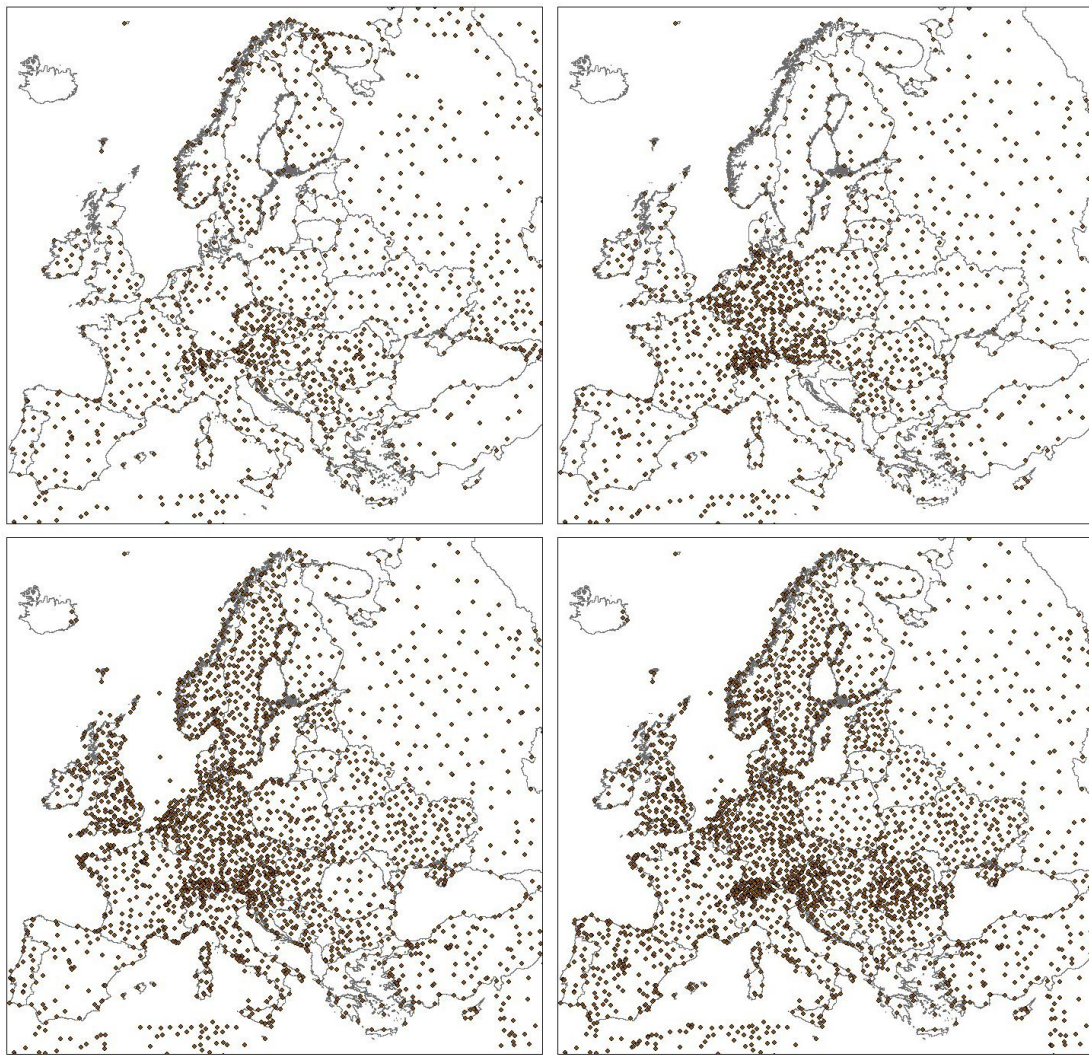


**Figure 10: Average annual cycle temperature difference for 1990-2011 between E-OBS and EFAS-Meteo databases. Panel (a) shows the minimum temperature difference and panel (b) shows the maximum temperature difference**

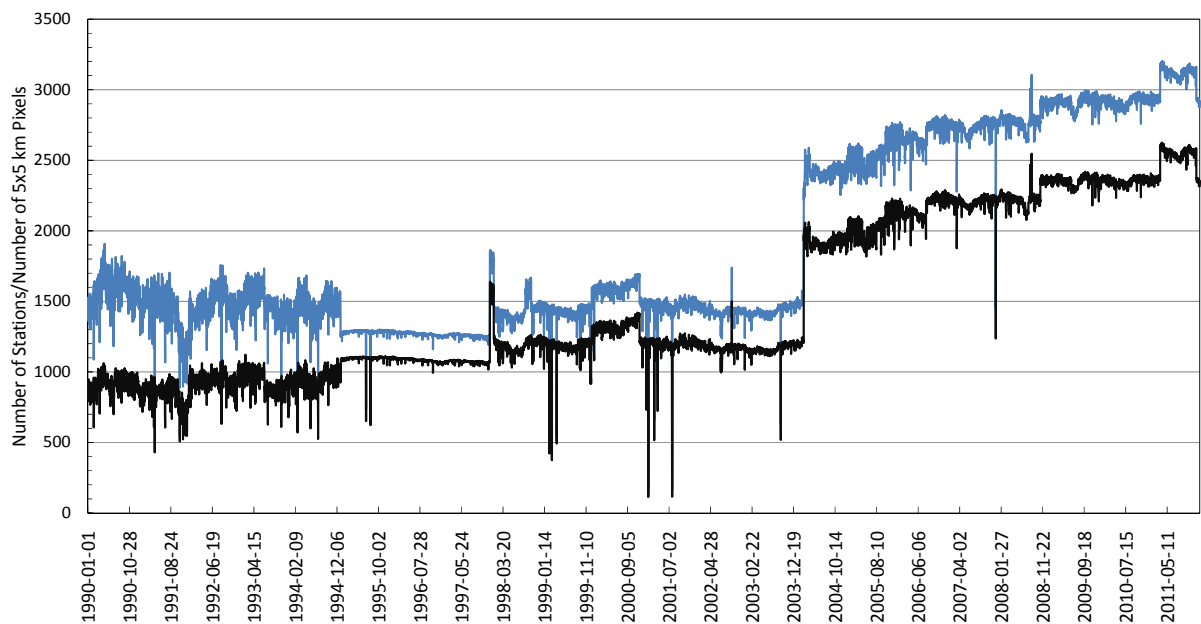
### 3.3 Calculated Radiation

Because of the scarcity of direct measurements of global radiation, the radiation is estimated using a staggered approach. When sunshine duration at a station is available global radiation is calculated using the Ångström-Prescott formula (Ångström, 1924; Prescott, 1940). When sunshine duration is not available but minimum and maximum temperature and cloud cover are known, the extended Hargreaves formula is applied (Supit, 1994; Supit and Kappel, 1998). Finally, when only minimum and maximum temperatures are known, the Hargreaves equation is applied for stations where the necessary constants have been calibrated before (Hargreaves, 1985). For a more detailed explanation on the calculated radiation the reader is referred to Baruth et al. (2007). Figure 11 exhibits the spatial distribution of the available stations for the selected periods 1990 to 2011. Figure 12 shows the time series of available stations for the period 1990-2011. The calculated radiation was derived using the JRC MARS database only.





**Figure 11 Spatial distribution of available calculated radiation observations for 31 Dec 1990 (upper left), 31 Dec 1995 (upper right), 31 Dec 2005 (lower left), and 31 Dec 2011 (lower right)**

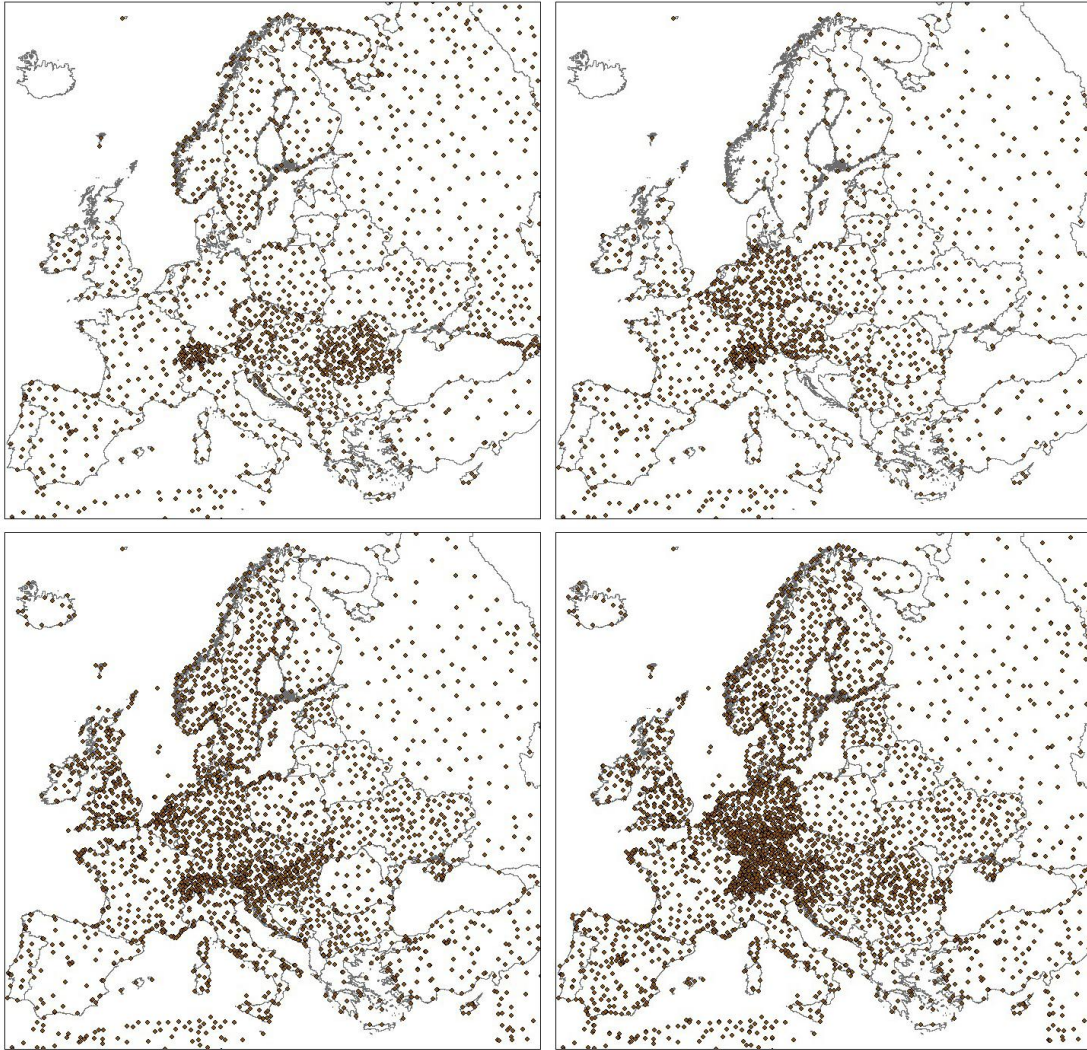


**Figure 12 Time series of available calculated incoming daily solar radiation time series for the JRC MARS source. The black time series shows the number of 5x5 km<sup>2</sup> pixels in the PCRaster map with one or more observations**



### 3.4 Wind speed

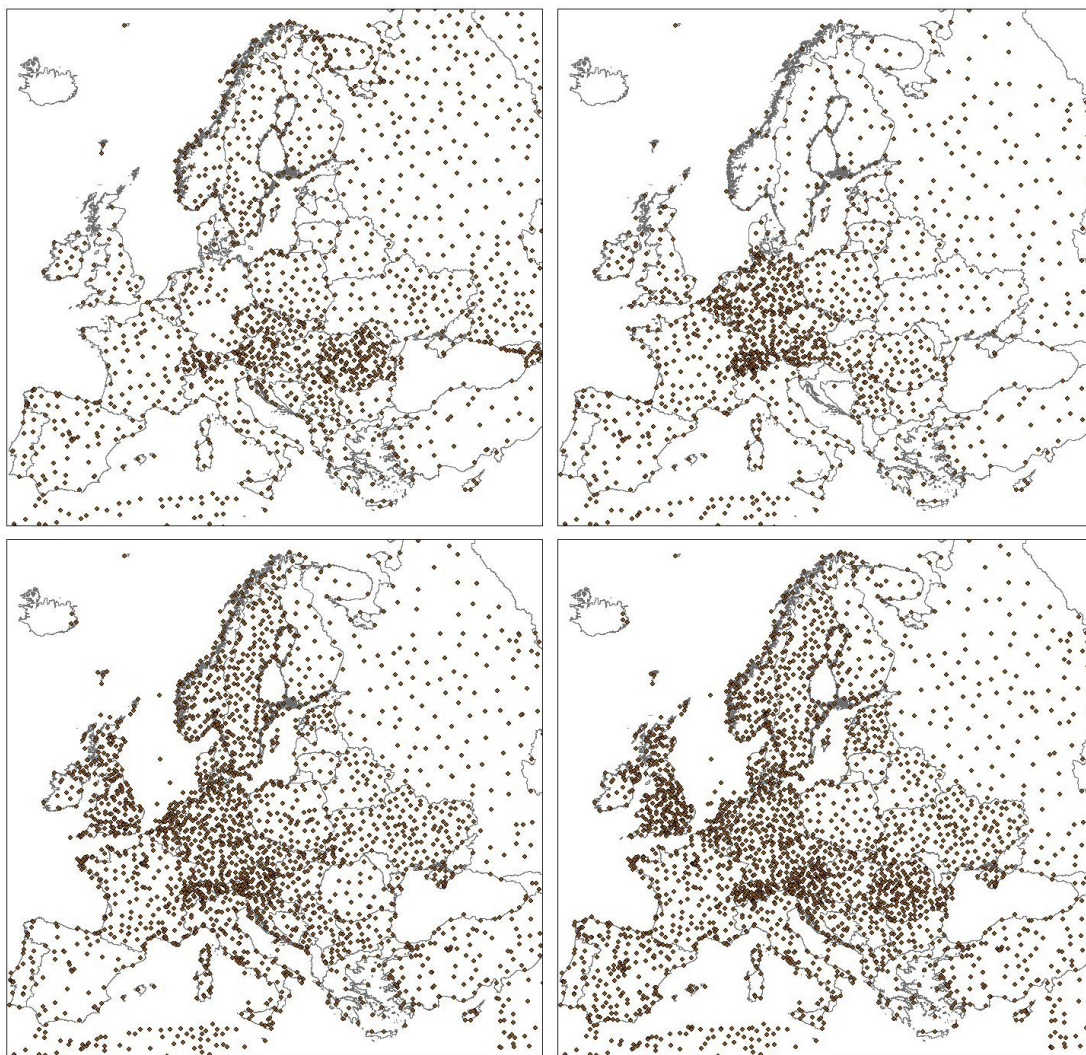
Mean daily wind speed is provided at 10 metres (m/s) and is calculated as the average speed for all available observations between 0-24 UTC. For calculations of evapotranspiration, this speed is transformed to a 2m speed (see Equation (19)). On average approximately 1500 5x5 km<sup>2</sup> pixels in the PCRaster maps have at least one observation for the time period between 1990 and 2003. Since about 2004 on average more than 2000 5x5 km<sup>2</sup> pixels have at least one or more observations. The spatial distribution is shown in Figure 13. The interpolation of wind fields is not without limitations. For instance, the interpolated fields are not robust enough to capture local effects such as orographic channelling which require a numerical wind model for better representation. Nevertheless the interpolated fields offer a better alternative to the assumption of a constant wind speed (less spatially varying wind speed) which is made when estimating evapotranspiration with limited data.



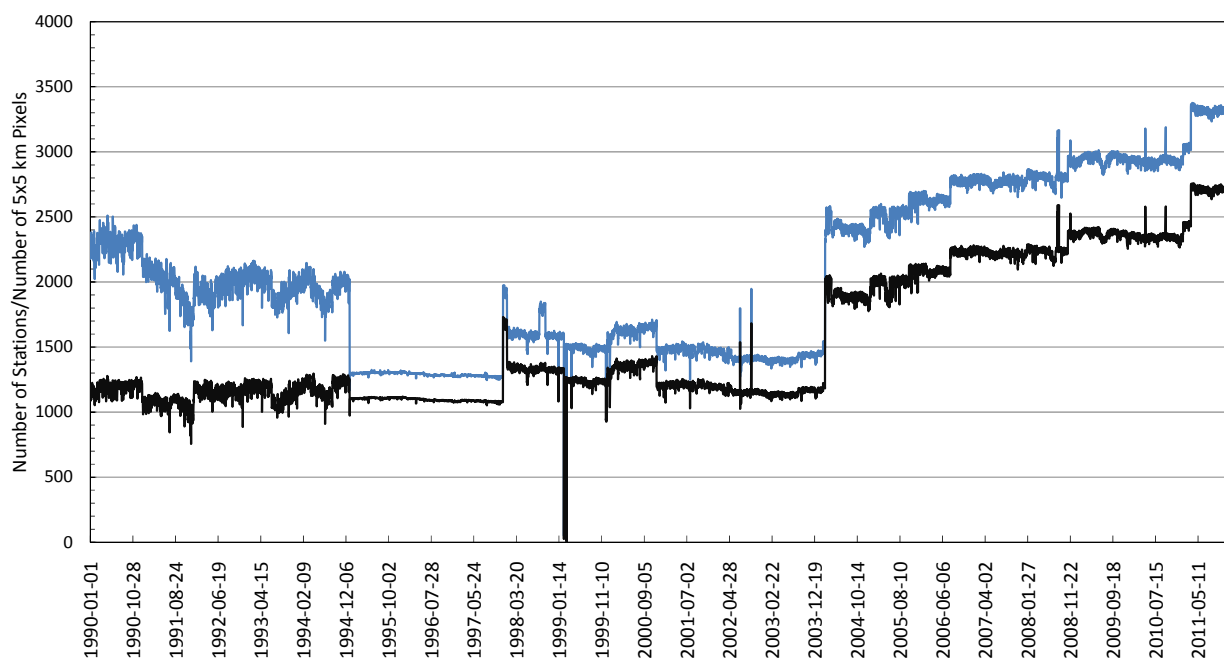
**Figure 13 Spatial distribution of available wind speed observations for 31 Dec 1990 (upper left), 31 Dec 1995 (upper right), 31 Dec 2005 (lower left), and 31 Dec 2011 (lower right)**

### 3.5 Vapour pressure

The vapour pressure, like the radiation, is extracted solely from the JRC MARS database. The extraction process incorporated an altitude correction similar to temperature but assuming a drop of 2.5% for every 100m rise in altitude. The available station distribution is shown in Figure 14. Figure 15 displays the time series of the available stations. It is shown that the station distribution is reasonably high and evenly distributed with at least 1000 5x5km<sup>2</sup> pixels containing at least 1 vapour pressure station for the period 1990-2011.



**Figure 14 Spatial distribution of available vapour pressure observations for 31 Dec 1990 (upper left), 31 Dec 1995 (upper right), 31 Dec 2005 (lower left), and 31 Dec 2011 (lower right)**



**Figure 15 Time series of available vapour pressure extracted from the JRC MARS database (blue line). The black line represents the number of 5x5 km pixels with one or more observations**



### 3.6 Evapotranspiration

Evaporation and water uptake and subsequent transpiration by vegetation are important components of the water balance. There are three climatic variables (mean daily temperature, wind speed, calculated radiation) required for the total estimation of evapo(trans)piration which have been extracted from the databases as described in the previous sections. First, a 'potential reference' evapotranspiration rate,  $ET$  is calculated. This is the evapotranspiration rate from a hypothetical reference vegetation with specific characteristics with unlimited availability of water (Allen et al., 1998). Similarly, a potential soil evaporation rate,  $ES$ , and the potential evaporation of an open water surface,  $EO$ , are calculated. Note that  $ET$ ,  $ES$  and  $EO$  are strictly climatic variables; they are not influenced by any land use or soil properties. In reality, the potential evapotranspiration can be either higher or lower than  $ET$  due to differences in vegetation characteristics, aerodynamic resistance and surface reflectivity (albedo).

The approach is based on the Penman-Monteith equation, and the procedure followed is mostly based on earlier work described by Supit et al. (1994) and Supit & Van Der Goot (2003). The calculation of potential evapo(transpi)ration is complicated somewhat by the fact that not all the necessary data is available for each time instant. Additionally the different datasets that are available are quite heterogeneous. For instance, incoming solar radiation can be estimated from sunshine duration or cloud cover data. Some data suppliers do not offer this kind of information, but provide pre-calculated grids of components of the radiation balance instead. Vapour pressure is sometimes substituted by dew point temperature. However, various equations have been proposed to handle the heterogeneity of such data. The sections that follow describe briefly how this is implemented. For a more detailed description of the calculation of evapotranspiration the reader is referred to van der Knijf (2008).

Reference values for potential evapotranspiration and evaporation are estimated using the Penman-Monteith equation (Supit *et al.*, 1994, Supit & Van Der Goot, 2003):

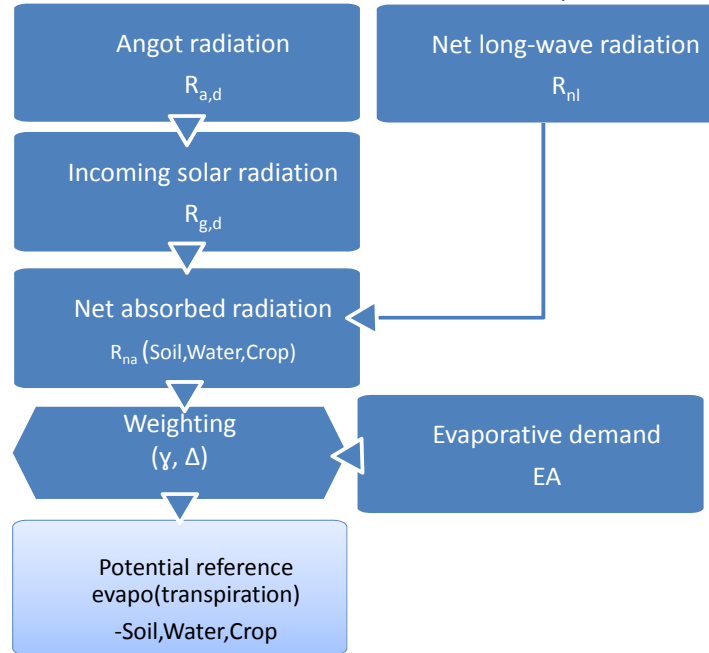
$$ET = \frac{\Delta R_{na} + \gamma EA}{\Delta + \gamma} \quad (1)$$

where:

$ET$	: Potential evapotranspiration rate from a closed vegetation canopy (reference crop) [mm day <sup>-1</sup> ]
$R_{na}$	: Net absorbed radiation [mm day <sup>-1</sup> ]
$EA$	: Evaporative demand of the atmosphere [mm day <sup>-1</sup> ]
$\Delta$	: Slope of the saturation vapour pressure curve [mbar °C <sup>-1</sup> ]
$\gamma$	: Psychrometric constant [mbar °C <sup>-1</sup> ]

The same equation is also used to estimate potential evaporation from a water surface and the evaporation from a (wet) bare soil surface (by using different values for the absorbed radiation term,  $R_{na}$ ). The procedure to calculate potential evapo(transpi)ration is summarized in Figure 16.

Table 4 lists the properties of the reference surfaces that are used in the computation of  $ET$ ,  $ES$  and  $EO$ , respectively.



**Figure 16 Procedure to calculate potential evapotranspiration**

**Table 4 Properties of reference surfaces for  $ET$ ,  $ES$  and  $EO$**

	$\alpha$ (surface albedo)	$f_e$ (empirical constant in evaporative demand equation)
<b><math>ET</math></b>	0.23	1

<b>ES</b>	0.15	0.75
<b>EO</b>	0.05	0.5

Calculating the net absorbed radiation term involves the following steps:

1. Calculate the daily extra-terrestrial radiation (Angot radiation)
2. From the Angot radiation, calculate the daily incoming solar radiation (using information on the daily number of sunshine hours or cloud cover, if available)
3. Calculate the daily net long-wave radiation (based on meteorological conditions)
4. Calculate the net absorbed radiation

The daily extra-terrestrial radiation is the product of the solar constant at the top of the atmosphere and the integral of the solar height over the day:

$$R_{a,d} = S_{c,d} \int \sin \beta dt_h \quad (2)$$

where:

$$\begin{aligned} R_{a,d} &: \text{Daily extra-terrestrial radiation [J m}^{-2}\text{day}^{-1}] \\ S_{c,d} &: \text{Solar constant at the top of the atmosphere [J m}^{-2}\text{s}^{-1}] \\ \int \sin \beta dt_h &: \text{Integral of the solar height over the day [s]} \end{aligned}$$

The solar constant on a given day is calculated as:

$$S_{c,d} = S_c (1 + 0.033 \cos[\frac{360 t_d}{365}]) \quad (3)$$

where:

$$\begin{aligned} S_c &: \text{Average solar radiation at the top of the atmosphere [J m}^{-2}\text{s}^{-1}] (= 1370 \text{ J m}^{-2}\text{s}^{-1}) \\ S_{c,d} &: \text{Solar constant at the top of the atmosphere [J m}^{-2}\text{s}^{-1}] \\ t_d &: \text{Calendar day number (1}^{\text{st}} \text{ of January is 1, etcetera) [-]} \end{aligned}$$

The calendar day number is always a number between 1 and 365.25 (taking into account leap years, a year has on average 365.25 days). The integral of the solar height equals:

$$\int \sin \beta dt_h = 3600 \{ L_d \cdot \sin \delta \cdot \sin \lambda + \frac{24}{\pi} \cdot \cos \delta \cdot \cos \lambda \cdot \sqrt{1 - (\tan \delta \cdot \tan \lambda)^2} \} \quad (4)$$

where:

$$\begin{aligned} L_d &: \text{Astronomical day length [h]} \\ \delta &: \text{Solar declination [}^\circ\text{]} \\ \lambda &: \text{Latitude [}^\circ\text{]} \end{aligned}$$

The solar declination is a simple function of the calendar day number ( $t_d$ ):

$$\delta = -23.45 \cdot \cos[\frac{360 (t_d + 10)}{365}] \quad (5)$$

Day length is given by:

$$L_d = \begin{cases} 12 + \frac{24}{180} a \sin(B_{ld}) & B_{ld} \geq 0 \\ 12 + \frac{24}{180} [a \sin(B_{ld}) - 360] & B_{ld} < 0 \end{cases} \quad (6)$$

with:

$$B_{ld} = \frac{-\sin(\frac{PD}{\pi}) + \sin \delta \cdot \sin \lambda}{\cos \delta \cdot \cos \lambda} \quad (7)$$

where  $PD$  is a correction constant (-2.65).

The incoming solar radiation is estimated using the Hargreaves equation as:

$$R_{g,d} = R_{a,d} A_h \sqrt{(T_{\max} - T_{\min})} + B_h \quad (8)$$

where:

$$\begin{aligned} A_h &: \text{Empirical constant [}^\circ\text{C}^{-0.5}] \\ B_h &: \text{Empirical constant [J m}^{-2}\text{d}^{-1}] \end{aligned}$$

The following equation is used to calculate the net long-wave radiation<sup>1</sup> (Maidment, 1993):

$$R_{nl} = f \varepsilon' \sigma (T_{av} + 273)^4 \quad (9)$$

where:

$R_{nl}$	: Net long-wave radiation [ $\text{J m}^{-2} \text{ day}^{-1}$ ]
$\sigma$	: Stefan Boltzmann constant ( $4.90 \times 10^{-3}$ ) [ $\text{J m}^{-2} \text{ K}^{-4} \text{ day}^{-1}$ ]
$f$	: Adjustment factor for cloud cover
$\varepsilon'$	: Net emissivity between the atmosphere and the ground

The net emissivity is calculated as:

$$\varepsilon' = 0.56 - 0.079 \sqrt{e_a} \quad (10)$$

with:

$e_a$	: Actual vapour pressure [mbar]
-------	---------------------------------

Two approaches were investigated for the cloud factor. Brunt (1932) and Allen (1994) proposed different methods for estimating the cloud factor:

The cloud cover factor according to Brunt (1932) is:

$$f = (B_e + B_f \frac{n}{L_d}) \quad (11)$$

where:

$e_a$	: Actual vapour pressure [mbar]
$B_e, B_f$	: Constants according to Brunt (1932) ( $B_e=0.2, B_f=0.9$ ) [-] or latitude varying factors.

If no information on the number of bright sunshine hours is available, the relative sunshine duration term is estimated using the Ångström equation:

$$\frac{n}{L_d} = \frac{(R_{g,d} / R_{a,d}) - A_a}{B_a} \quad (12)$$

where  $A_a$  and  $B_a$  are the empirical Ångström constants.

The cloud cover according to Allen et al. (1994) is calculated as:

$$f = (1.8 \frac{R_{g,d}}{R_{so}} - 0.35) \quad (13)$$

where:

$R_{so}$  is the clear sky total global solar radiation at the Earth's surface in  $\text{MJm}^{-2} \text{ d}^{-1}$ :

$$R_{so} = R_a (0.75 + 2.00 \times 10^{-5} E_L)$$

where:  $E_L$  is the elevation in meters. The coefficient (1.8) in Equation (13) was adjusted from the original coefficient (1.35) suggested by Allen et al. (1994) to account for a higher cloudiness in Europe compared to areas in USA where the coefficients were locally fitted (Dorrenbos and Pruitt, 1977). Also  $f$  cannot be negative or greater than 1. Thus limits of 0.05 and 1 were taken for such cases.

A comparison of the two cloud factors found that the evapotranspiration totals were not significant. Moreover the two equations can be recalibrated to fit local conditions better. However, the Allen et al. (1994) approach explicitly accounts for elevations up to 6000m, which is robust for the highland as well as lowland areas. Hence, the Allen et al. (1994) approach was selected for estimating the cloud factor.

Finally, the net absorbed radiation [ $\text{mm day}^{-1}$ ] is calculated as:

$$R_{na} = \frac{(1 - \alpha) R_{gd} - R_{nl}}{L} \quad (14)$$

where  $\alpha$  is the albedo (reflection coefficient) of the surface, and  $L$  is the latent heat of vaporization [ $\text{MJ kg}^{-1}$ ]:

$$L = 2.501 - 2.361 \cdot 10^{-3} \cdot T_{av} \quad (15)$$

The net absorbed radiation is calculated for three cases: for a reference vegetation canopy (using  $\alpha=0.23$ ), a bare soil surface ( $\alpha=0.15$ ), and an open water surface ( $\alpha=0.05$ ) The evaporative demand of the atmosphere is calculated as:

<sup>1</sup> Note that this term is mistakenly called 'net outgoing longwave radiation' in the WODOST/CGMS documentation (Supit et. al., 2003), whereas it is in fact the net longwave radiation

$$EA = 0.26[e_s - e_a][f_c + BU \cdot u(2)] \quad (16)$$

where:

$EA$	: Evaporative demand [mm day <sup>-1</sup> ]
$e_s$	: Saturated vapour pressure [mbar]
$e_a$	: Actual vapour pressure [mbar]
$f_c$	: Empirical constant [-]
$BU$	: Coefficient in wind function [-]
$u(2)$	: Mean wind speed at 2 m height [m s <sup>-1</sup> ]

Saturated vapour pressure is calculated as a function of mean daily air temperature:

$$e_s = 6.10588 \cdot e^{\frac{17.32491 \cdot T_{av}}{T_{av} + 238.102}} \quad (17)$$

The coefficient in the wind function,  $BU$ , is also temperature dependent:

$$BU = \max \left[ 0.54 + 0.35 \frac{\Delta T - 12}{4}, 0.54 \right] \quad (18)$$

Here,  $\Delta t$  is the difference between the daily maximum and minimum temperature. Equation (18) implies that  $BU$  has a fixed value of 0.54 if  $\Delta t$  is less than 12°C.

Since wind speed is usually measured at a height of 10 m, the following correction is made (Maidment (1993), p. 4.36):

$$u(2) = 0.749 \cdot u(10) \quad (19)$$

where  $u(10)$  is the measured wind speed at 10 m height [m s<sup>-1</sup>].

Similar to the calculation of the net absorbed radiation, the evaporative demand is calculated for three cases: for a reference vegetation canopy (using  $f_c = 1.0$ ), a bare soil surface ( $f_c = 0.75$ ), and an open water surface ( $f_c = 0.5$ ).

The psychrometric constant at sea level can be calculated as:

$$\gamma_0 = 0.00163 \frac{P_0}{L} \quad (20)$$

where:

$\gamma_0$	: Psychrometric constant at sea level (about 0.67) [mbar °C <sup>-1</sup> ]
$P_0$	: Atmospheric pressure at sea level [mbar]
$L$	: Latent heat of vaporization [MJ kg <sup>-1</sup> ]

Since the barometric pressure changes with altitude, so does the psychrometric constant. The following altitude correction is applied (Allen *et al.*, 1998):

$$\gamma(z) = \gamma_0 \left( \frac{293 - 0.0065 \cdot z}{293} \right)^{5.26} \quad (21)$$

where:

$\gamma(z)$	: Psychrometric constant at altitude $z$ [mbar °C <sup>-1</sup> ]
$z$	: Altitude above sea level [m]

The slope of the saturation vapour pressure curve is calculated as follows:

$$\Delta = \frac{238.102 \cdot 17.32491 \cdot e_s}{(T + 238.102)^2} \quad (22)$$

where  $\Delta$  is in [mbar °C<sup>-1</sup>].

The potential evapo(transpi)ration can now be calculated for three reference surfaces (ET, ES, EW) according to Eq. (1)

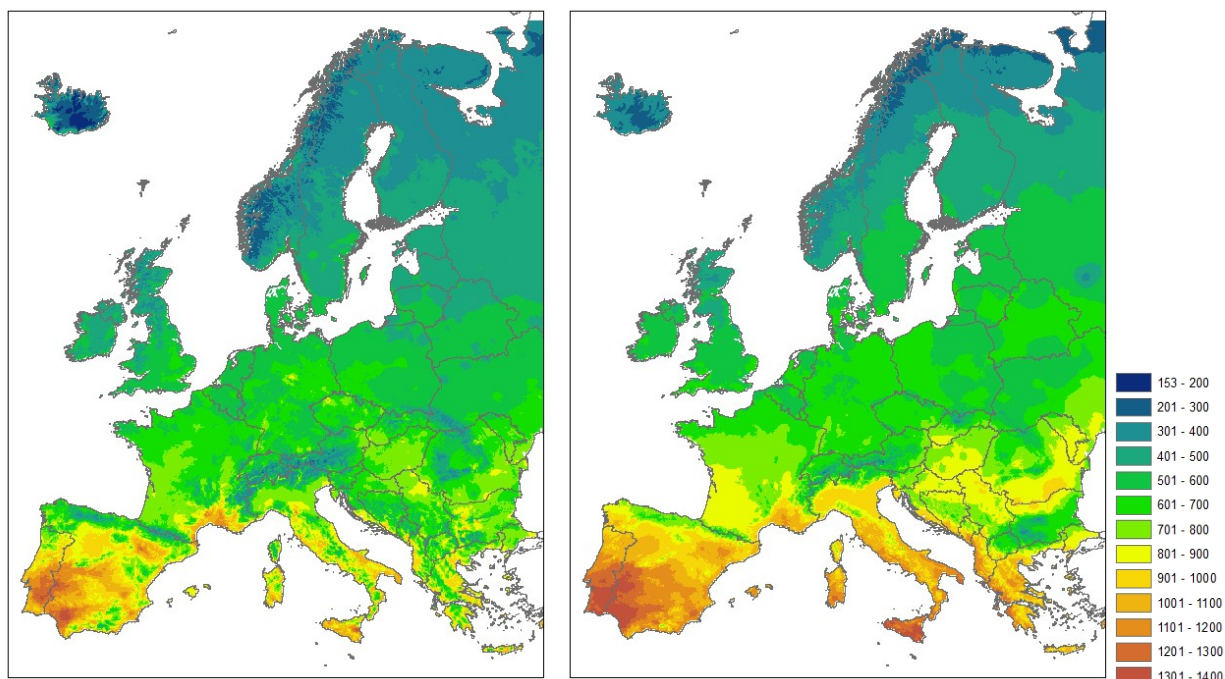
The Hargreaves equation for estimating potential evapotranspiration was developed for cases of limited data availability. While we have used for the EFAS-Meteo dataset a large dataset of different variables, it is useful to incorporate methods that require minimum data to provide an alternative in case of unforeseen limitations such as delays in data availability. In an operational sense, if data is not ready for calculating the data intensive penman monteith evapotranspiration, the Hargreaves approach would be a viable supplement.

The Hargreaves equation (Hargreaves *et al.*, 1985) is given as follows:

$$ET0 = 0.0135 R_{g,d} (T_{av} + 17.8) \quad (23)$$

$R_{g,d}$  can be either a direct measurement or derived using empirical relationships such as shown in Equation (8) However, when using the Hargreaves equation, it is only possible to calculate the potential evapotranspiration for a reference crop, as it implicitly accounts for the albedo of 0.23. Another approach could be to estimate the Penman Monteith ratios

between  $ET$  and  $ES$ , and  $ET$  and  $E0$  and use these as multiplication factors to the estimated Hargreaves  $ET$ . Figure 17 illustrates the annual average evapotranspiration for the Penman Monteith and Hargreaves equations. It is shown that the Hargreaves approach tends to be higher than the Penman Monteith especially in the southern regions. It is worthwhile to establish ways of applying bias corrections to the Hargreaves to minimize the differences with the Penman Monteith, which is considered a more physically based approach. Monthly Penman Monteith maps can be found in the appendix.



**Figure 17 Annual Penman Monteith Evapotranspiration (left) and Hargreaves evapotranspiration (right)**

## 4 Analysis of trends

This section is aimed at detecting evidence of trends in the EFAS-Meteo dataset at a pan-European scale. Although the time series is quite short, the findings are valuable for a number of reasons. First, the high resolution of the dataset is vital for analysing extremes especially regarding rainfall. High resolution datasets have been found to better represent extreme rainfall especially in areas of complex topography (Zolina et al, 2013). Other existing pan-European datasets are at coarser scales, which include, in some cases, considerably less observations. Second, it is clear that given the existing disparities in the dataset, spurious trends are expected. Thus trends are useful as diagnostics for data quality. To check improvements in data quality, the analysis can be repeated to ensure that spurious trends are diminished. Third, regional and temporal differences are highlighted, which could be an indication of changes in physical phenomenon that once established, helps to clarify the underlying drivers of the trends. Fourth, through comparative analysis with other data archives interpretation of trends would be enhanced. However, it is clear that with such short time series, making strong statements is not recommended. The findings only provide a small scale view of the trends and different interpretations could arise with large scale views.

### 4.1 Trend analysis procedure

Trends are analyzed using a generalized least squares approach that is based on restricted maximum likelihood (REML) assuming residuals follow a first order autoregressive process (AR1). This ensures that the temporal autocorrelation is accounted for (Foster and Rahmstorf, 2011) to allow for more accurate analysis of trend significance. Usually if a significant autocorrelation exists, the variance of the data is underestimated and this leads to a narrow confidence interval. This means the presence of significant trends is more likely. The REML analysis is carried out for each grid point and at a seasonal and annual scale. The seasonal extremes are categorized annually so as to assess the trends. Trend slopes and statistical p-values are calculated. For a 95 percent confidence interval, a p-value of 0.05 represents an important threshold. When the p-value is lower, the trend is statistically significant when it is higher it is insignificant.

#### 4.1.1 Annual temperature trends

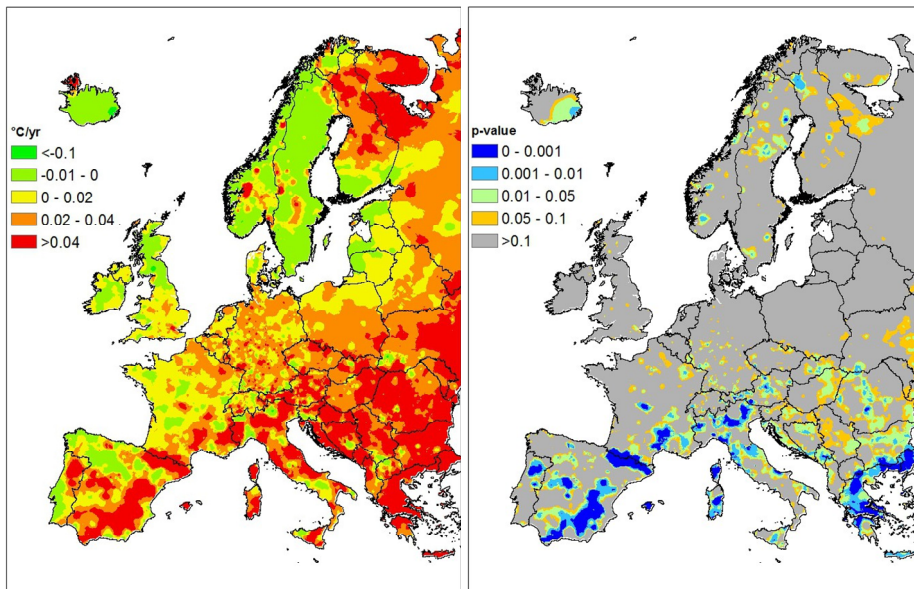
Average temperature trends are calculated for each grid point and corresponding slope (trend) and p-value noted. Results for all the grid points are shown in Figure 18. The trends are fitted for period 1990-2012 for average temperature (average of Tmin and Tmax). It is clear that southern Europe has higher warming rates with significant warming trends for Spain, southern France, northern Italy and Greece. Trends trend higher than +0.04°C/year (red zones) is observed. High warming rates (>+0.04°C) are also observed for parts of northern Europe (northwest Russia and Finland). On average, the rate of warming was +0.2 °C/decade (+0.02 °C/year). A 4°C world by 2100 (trend of +0.04°C/year) would greatly disrupt water, ecosystems, food, coasts and human health (New et al., 2011).

#### 4.1.2 Seasonal temperature trends

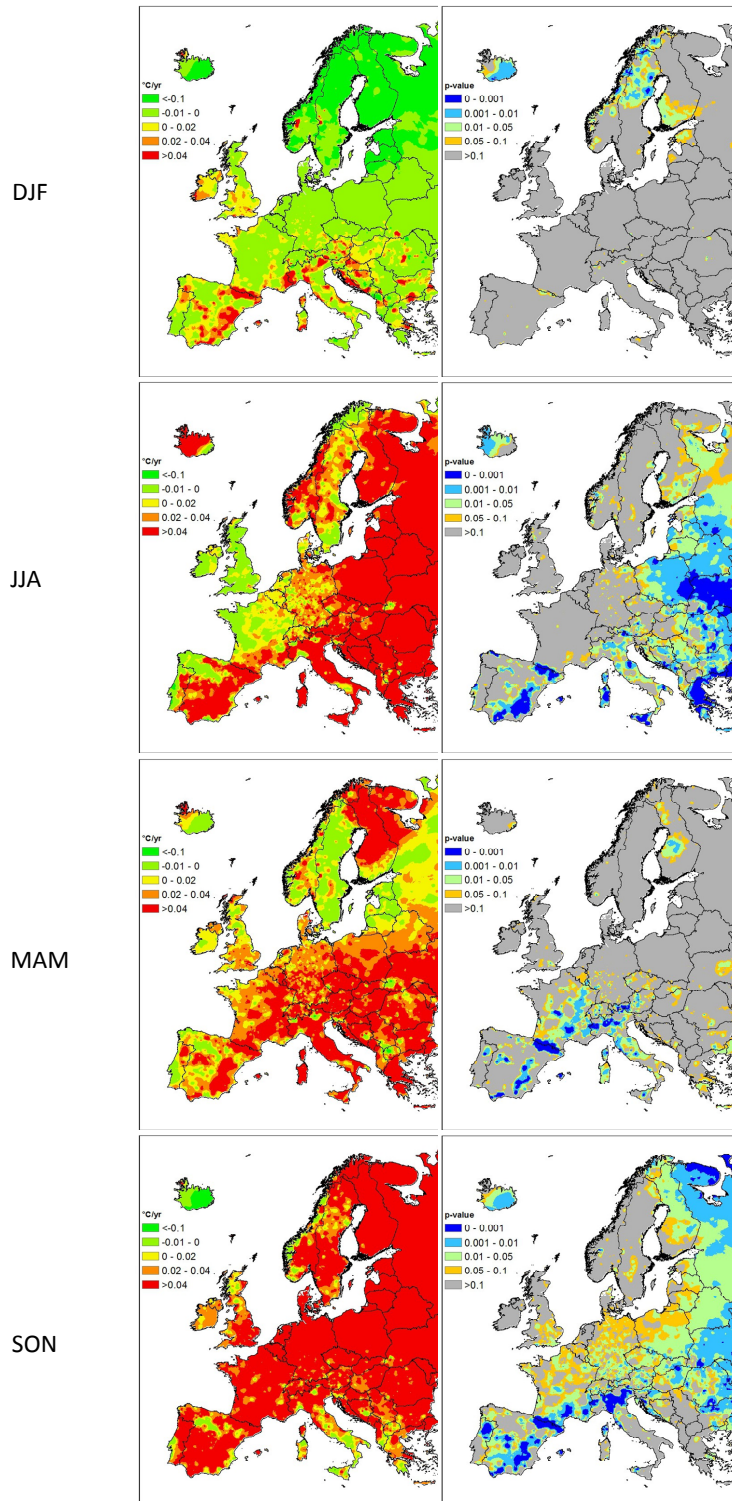
In contrast to other seasons, winter showed dominant cooling (average of -0.05°C/year) during the period 1990-2012 (Figure 19). Downward trends were observed in the north and warming trends in parts of southern Europe. Downward trends were significant for northern areas of Norway and Sweden. Summer was predominantly warm (average of +0.05°C/year) with most areas in eastern and southern Europe showing significant warming trends. Autumn had strongest warming trends (average of +0.07 °C/year) of all the seasons. Significant trends were observed for north western Russia and southern Europe (Spain, France, and Italy). Spring experienced warming trends (average +0.03°C/year) in most of Europe, but trends were significant in southern Europe (Spain, France and Italy).

These results are in line with various trend records from different datasets (Jones et al., 2013; Girvetz et. al., 2009). It is important to emphasize that other comparable pan-European studies have used longer time series at coarser scales for analysing trends, which means there are differences in seasonal and annual trend characteristics. Jones et al. (2013), found that over the period 1850-2012, all seasons show a general upwards trend from the 1970s, but the last 10 years have experienced a general levelling off in temperature anomalies in the Northern Hemisphere. However, this levelling off occurs in a period with some of the highest recorded temperatures. Nine warmest years occurred since 2001; only 2008 and 2011 were not in the ten warmest years on record.





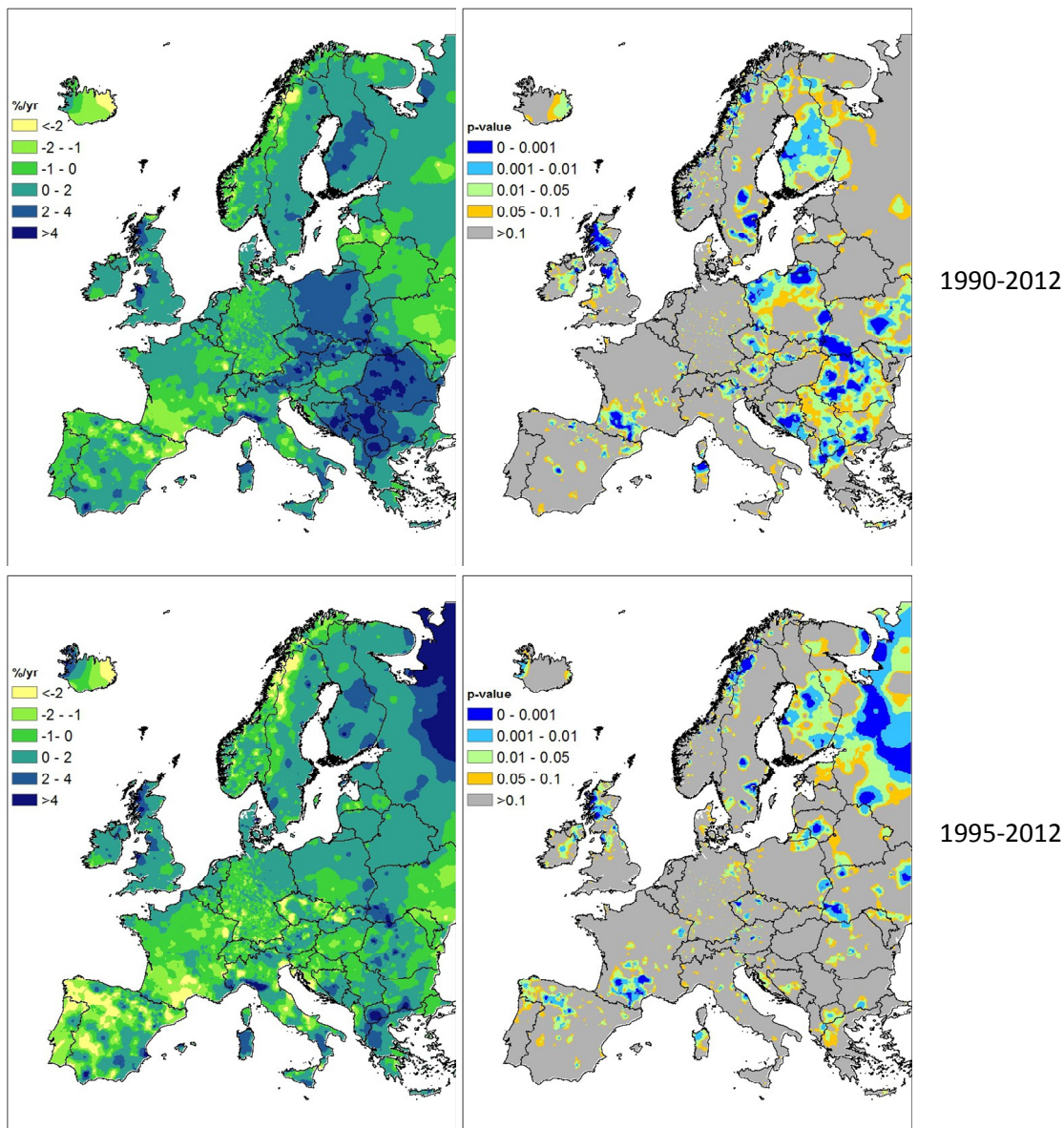
**Figure 18 Mean annual temperature trends (left) significance p-values (right) for 1990-2012**



**Figure 19 Seasonal temperature trends (left) and significance p-values (right) for 1990-2012**

### 4.1.3 Annual precipitation trends

Annual trends are estimated for two time windows (Figure 20) so as to show more accurate trends where data quality is a known limitation. As previously discussed, eastern Europe showed evidence of disparities for some regions. For such regions, trends are better assessed using the time window 1995-2012. When assessed for the entire period 1990-2012, eastern Europe shows spurious significant trends. The eastern significant zone is evident in Figure . Western and northern Europe show increasing precipitation trends but they are not significant (1990-2012). Based on the 2 temporal windows, some regions show persistent trends. Significant increasing trends are persistent for northern UK (Scotland) and significant decreasing trends are observed for southern France. Evidence of decreasing trends in southern France are in line with other trend studies( Philandras et al.;2011;Vidal et al.,2010). Increasing trends in Scotland have been reported in Simpson and Jones (2013). Other regions do not show persistent trends for the two time windows. It is important to point out that the Russian region should be ignored in the analysis as precipitation data in this region was not reliable for analysis.



**Figure 20 Annual precipitation trends (right) for 1990-2012 (top) and 1995-2012 (bottom) along with p-values (right)**

#### 4.1.4 Seasonal precipitation trends in extremes

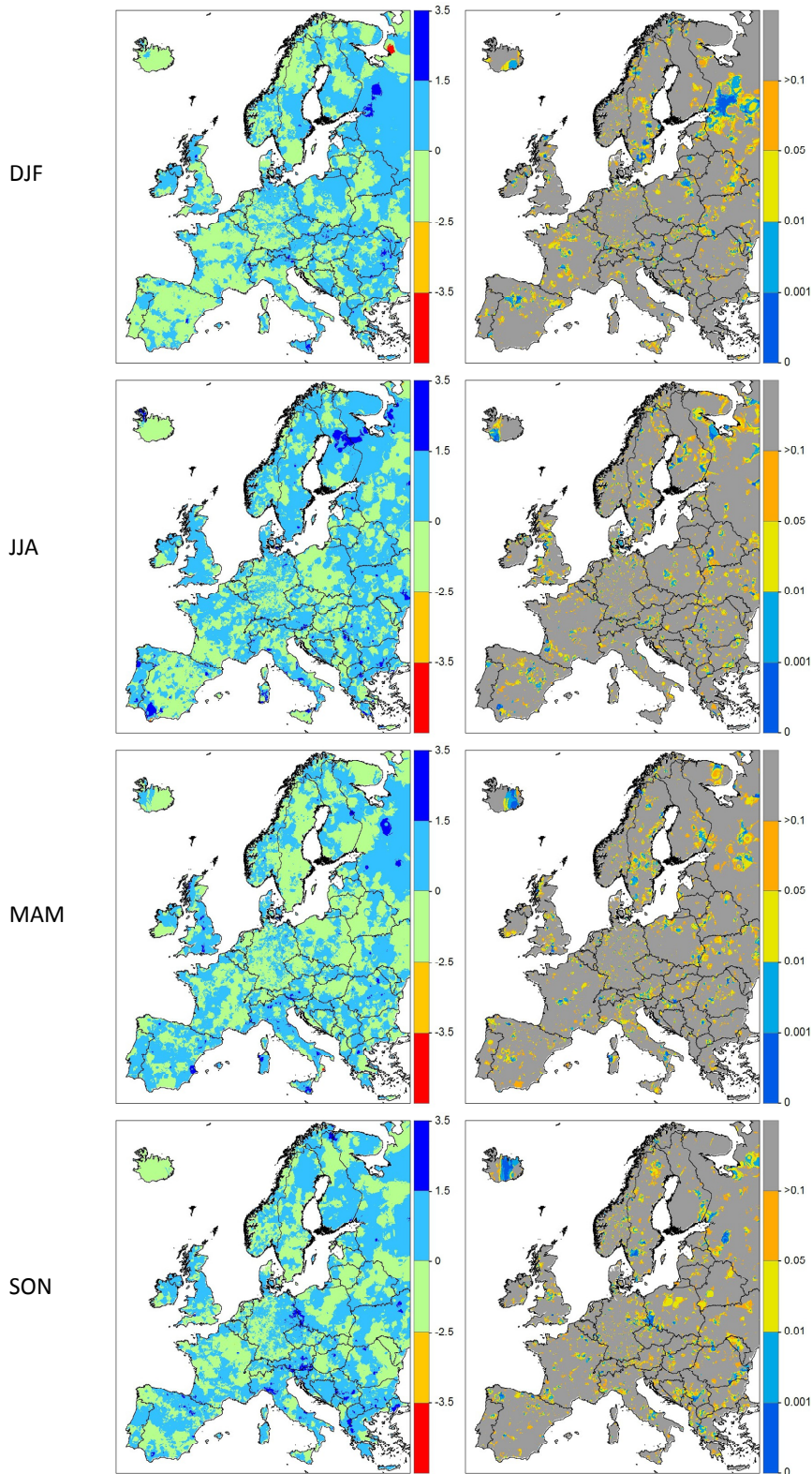
The non-parametric approach applied for trends spanning the period 1995-2012 is taken from the study of Osborn and Hulme (2002). In that study, category 10 events are defined as those heaviest events that contribute 10% of the total aggregated rainfall for the period. However, the total is calculated on a monthly basis and for days that are considered wet, that is, days greater than a certain threshold (say 0.4mm). In most cases, 0.5-1.5% of the heaviest events are sufficient to contribute to 10% of the total rainfall. To perform this analysis, all the wet days in a particular month (18 months for 1995-2012) period, are sorted in ascending order and the top events that contribute to 10 % of the total rainfall amount are determined. Because these events are distributed in time, the percentage contribution by year can be estimated. This means that it is likely that there are years where the contribution of precipitation extremes will be zero. After determining the contribution for each month, seasonal averages are estimated. For example, for winter, average percentage contribution is derived from the contributions of January, February, and December. Here, winter timestamp is defined by the year of January. A trend is then fitted to the series and its significance determined. The trend is taken as an indicator of increase in extreme precipitation intensity.

This approach varies from other non-parametric approaches which rely on globally fixed thresholds such as events greater than 20 mm to classify extremes. For globally fixed thresholds, there is no guarantee that there will be enough events to perform the trend analysis. Moreover, using such thresholds one assumes that the extreme event definition is not spatially varying. Additionally, performing this analysis for each of the 12 months, means that when assessing seasonal or area averages, changes are not overly influenced by months with significant changes or biased by shifts in seasonality.

During winter, most parts of western Europe show downward trends in the percentage contribution from extreme events with the exception of western and northern parts of UK (Figure 21). However, this trend is not significant for the period 1995-2012. Increasing non-significant trends dominate in eastern Europe.

Summer shows upward trends for many areas in Europe with the exception of Spain, which shows similar drying trends with winter. Regions of Hungary, Slovakia and Poland show decreasing trends which is quite different from the increasing trends observed for surrounding areas. However, the statistical significance is weak for many regions. Considering the two main seasons, that is winter and summer, a distinct pattern emerges showing upward trends in eastern Europe and UK. Spain shows consistent downward trends. Other regions show differences in directions of change for winter and summer. Decreasing trends are apparent for northern Europe during spring, with the exception of UK and Norway which show increasing trends. Increasing trends dominate southern Europe. Significance of trends is, however, weak for many regions. On average, most parts of Europe show increasing trends in autumn. Decreasing trends dominate in Spain, Netherlands, Romania, Hungary, Moldova, Latvia, Lithuania and Belarus. Significance of trends is, however, weak for many regions.





**Figure 21 Seasonal trends in extreme precipitation contribution (%/year) for 1995-2012 (left) along with p-values (right)**

The aforementioned trend findings are somewhat similar to other studies related to precipitation extremes. The consistency is not obvious to compare due to differences in methodologies, different temporal windows, and spatial scales. Nevertheless, it is interesting to note the similarities and differences or shifts in trends. Dominating negative summer and winter trends, and positive trends during spring and autumn for Spain appear in line with other studies (Rio et al., 2011; De Luis et al., 2009). Evidence of recent dry winters and wet summers for UK is not in line with prior studies (Maraun et al., 2008; Osborn et al., 2000). These differences mainly arise from different time windows. Interestingly, Simpson and Jones (2013) found that recent trends in UK were characterized by successive dry climates in winter with successive wetter summers. This is consistent with the results from EFAS-Meteo. Increasing trends in summer contribution to extremes can be explained by wet summers starting in 2007 (Simpson and Jones, 2013; Coumou and Rahmstorf, 2012). This highlights the limitation of analysing trends over short time windows or

extrapolation of trends from past records. If prior observations are extended to most recent periods, the reported negative decreasing trends during summer, and positive trends during winter would be offset.

## 5 Summary and conclusions

The EFAS-Meteo dataset has been in preparation for over a decade. Since then the dataset has grown to include more providers following bilateral agreements made with partners. The dataset now covers the entire European region. With a compendium of providers, the challenges of harmonising the dataset were not trivial. Given the diversity of providers following different data collection and storage procedures among other challenges, efforts were made to harmonise the data. The overriding principle was to follow procedures that would 1) ensure quality, and 2) lead to a database system that would be operationally feasible. The dataset revealed that there are some shortcomings in some regions of Europe prior to 1990, and some regions prior to 1995. Notwithstanding the few regional shortcomings, the utility of the dataset has demonstrated how essential it is to have such a high resolution dataset. The high resolution particularly enables a more accurate representation of extreme characteristics. What's more, the dataset forms an important kernel of the operational EFAS setup, and will be an important reference for other hydro-meteorological studies. It is expected that the database will continue to improve as more data archives and technologies become available. This will especially be valuable for regions where the dataset has shown shortcomings. Additionally, the procedures developed during the database creation process will enable more efficient updates.

## Acknowledgements

We would like to acknowledge the JRC MARS unit for its valuable support with the MARS database. Furthermore, we would like to acknowledge all the national hydrological and meteorological services listed in the Appendix for providing their data within the support of the European Floods Awareness System. Finally, we would like to acknowledge the reviewers of this document for providing valuable comments.

## References

- Allen, R. G., Smith, M., Perrier, A., and Pereira, L. S., (1994). An Update for the Definition of Reference Evapotranspiration, ICID Bulletin, ICID, 43(2), 1-34.
- Allen, R.G., Pereira, L.S., Raes, D., and Smith, M., (1998). Crop evapotranspiration guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56.
- Ångström, A., (1924). Solar and terrestrial radiation. Quarterly Journal of the Royal Meteorological Society, 50:121-125.
- Bartholmes, J. C., J. Thielen, M. H. Ramos, and S. Gentilini (2009). The European Flood Awareness System EFAS - Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts, Hydrology and Earth System Sciences, 13, 141-153.
- Baruth, G. Genovese, O. Leo Contributors: H. Boogard, J. A. te Roller, K. van Diepen, (2007). CGMS version 9.2 User manual and technical documentation.
- Brunt, D., (1932). Notes on radiation in the atmosphere. International Quarterly Journal of the Royal Meteorological Society 58, 389-420.
- Burrill, A., Vossen, P. (1992). Development of an operational agro-meteorological monitoring system. In: F. Toselli, J. Meyer-Roux (eds). Proceedings of conference on the application of remote sensing to agricultural statistics, 26-27 November 1991, Belgirate, Italy. EUR 14262 EN, Office for Official Publications of the EU, Luxembourg, p 357-360.
- Caesar, J., L. Alexander, and R. Vose (2006). Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set, J. Geophys. Res., 111, D05101, doi:10.1029/2005JD006280.
- Coumou, D., and Rahmstorf, S. (2012). A decade of weather extremes. Nature Climate Change, 2(7), 491-496.
- De Luis M., González-Hidalgo J., Longares, L., Stepánek, P. (2009). Seasonal precipitation trends in the Mediterranean Iberian Peninsula in second half of the 20th century. International Journal of Climatology 29(9): 1312-1323.
- Doorenbos, J., and Pruitt, W. O., (1997). Crop Water Requirements, FAO Irrigation and Drainage Paper 24, FAO, Roma.
- Foster, G. and Rahmstorf, S. (2011). Global temperature evolution 1979-2010. Environ. Res. Lett. 6, 044022.
- Frei, C., and C. Shar (1998). A precipitation climatology of the Alps from high-resolution rain-gauge observations, Int. J. Climatol., 18, 873- 900, doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255> 3.0.CO;2-9.
- Girvetz, E. H., C. Zganjar, G. T. Raber, E. P. Maurer, P. Kareiva, and Lawler, J. J. (2009). Applied climate-change analysis: The Climate Wizard tool. PloS One, 4, e8320, doi:10.1371/journal.pone.0008320.

- Goudriaan, J., (1977). Crop micrometeorology: a simulation study. Simulation monographs, Pudoc, Wageningen.
- Hargreaves, L. G., Hargreaves, G. H., and Riley, J. P., (1985). Irrigation Water Requirements for Senegal River Basin, J. Irrig. and Drain. Engrg., ASCE, 111(3), 265-275.
- Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New (2008). A European daily high resolution gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res., 113, D20119, doi:10.1029/2008JD010201.
- Hofstra, N., M. Haylock, M. New, P. Jones, and C. Frei (2008). Comparison of six methods for the interpolation of daily, European climate data, J. Geophys. Res., 113, D21110, doi:10.1029/2008JD010100.
- Jones, P.D., Parker, D.E., Osborn, T.J., Briffa, K.R. (2013). Global and hemispheric temperature anomalies—land and marine instrumental records, In Trends: A Compendium of Data on Global Change, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. doi: 10.3334/CDIAC/cli,002.
- Karssenbergh, D., C. G. Wesseling, and W. P. A. Van Deursen (1996). "PCRaster version 2 manual." The Netherlands, Department of Fysical Geography, University of Utrecht.
- Maidment, D.R. (editor), (1993). Handbook of Hydrology. McGraw-Hill.
- Maraun D., Osborn, T.J., and Gillett N.P. (2008). United Kingdom daily precipitation intensity: improved early data, error estimates and an update from 2000 to 2006. Int J Climatol 28(6):833–842. doi:10.1002/joc.1672
- MeteoConsult, 1991. AMDaC System Manual. MeteoConsult, Wageningen, The Netherlands, pp 40.
- New, M., Liverman, D., Schroder, H., and Anderson, K. (2011). Four degrees and beyond: the potential for a global temperature increase of four degrees and its implications. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369(1934), 6-19.
- Osborn T.J., Hulme, M. (2002). Evidence for trends in heavy rainfall events over the UK. Philosophical Transactions of the Royal Society London A360: 1313–1325.
- Osborn, T. J., Hulme, M., Jones, P. D. and Basnett, T. A. (2000). Observed trends in the daily intensity of United Kingdom precipitation. Int. J. Climatol., 20: 347–364.
- Philandras, C.M., Nastos, P.T., Kapsomenakis, J., Douvis, K.C., Tselioudis, G., and Zerefos, C.S. (2011). Long term precipitation trends and variability within the Mediterranean region Nat. Hazards Earth Syst. Sci., 11, pp. 3235–3250.
- Prescott, J.A., (1940). Evaporation from a water surface in relation to solar radiation. Transactions of the Royal Society of South Australia, 64:114-118.
- Río, S. del, Herrero, L., Fraile, R., and Penas, A. (2011). Spatial distribution of recent rainfall trends in Spain (1961–2006). International Journal of Climatology 31, no. 5: 656-667.
- Simpson, I. R., and Jones, P. D. (2013). Analysis of UK precipitation extremes derived from Met Office gridded data. International Journal of Climatology.
- Supit, I. & van Kappel, R.R., (1998). A simple method to estimate global radiation. Solar Energy 63 (3), 147-160.
- Supit, I. , van der Goot, E. (eds.), (2003). Updated System Description of the WOFOST Crop Growth Simulation Model as Implemented in the Crop Growth Monitoring System Applied by the European Commission, Treemail, Heelsum, The Netherlands, 120 pp.
- Supit, I., Hooijer, A.A. and van Diepen, C.A., (1994). System description of the WOFOST 6.0 crop simulation model implemented in CGMS, Volume 1: Theory and Algorithms. Joint Research Centre of the European Community, EUR 15956 EN.
- Thielen J., Antoniotti C., Nappo D., Gomes G., Sint H., Dusart J., Kalas M., Salamon P., Lorini V., Ntegeka V., Harris C., and Hohmann K. (2012). Handling floods of data - Data collection in support to the European Flood Awareness System. JRC Technical Report, EUR xxxx EN.
- Thielen-del Pozo, J., J. Bartholmes, M.-H. Ramos, and A. de Roo (2009). The European Flood Awareness System. Part 1: Concept and development, Hydrology and Earth System Sciences, 13, 125-140.

- Van der Kniff, J., (2008). LISVAP-Evaporation Pre-Processor for the LISFLOOD Water Balance and Flood Simulation Model, Revised User Manual. EUR 22639 EN/2, Office for Official Publications of the European Communities, Luxembourg, 31pp.
- Van der Knijff, J. Younis and A. de Roo (2010). LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation International Journal of Geographical Information Science; Vol. 24, No. 2, February 2010, 189–212.
- Vidal, J. P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J. M. (2010). A 50-year high-resolution atmospheric reanalysis over France with the Safran system. International Journal of Climatology, 30(11), 1627-1644.
- Zolina, O., Simmer, C., Kapala, A., Shabanov, P., Becker, P., Mächel, H., and Groisman, P. (2013). New view on precipitation variability and extremes in Central Europe from a German high resolution daily precipitation dataset: Results from the STAMMEX project. Bulletin of the American Meteorological Society.



## Technical Appendix

### Map format:

All maps are stored in PCRaster format and fulfil the following conditions:

rows:	810
columns:	680
cell size:	5000 m
x coordinate upper left corner:	-1.7e6
y coordinate upper left corner:	2.7e6

Details concerning the PCRaster map format can be found at <http://pcraster.geo.uu.nl/>. It is recommended, when transforming the maps into another raster format such as GeoTiff for example, to use the command line utility `gdal_translate` from the Geospatial Data Abstraction Library (GDAL - <http://www.gdal.org/>).

### Map file naming format:

The format of the file names of the PCRaster maps consists of total of 12 characters: 8 characters, a dot and 3 characters. The first two characters indicate the variable contained in the map, whereas the remaining characters are used to indicate the time step. A file name containing precipitation for time step 1 would look as follows: `pr000000.001`. A file name containing mean daily temperature for time step 9999 would look as follows: `ta000009.999`. etc. The list of abbreviations for the meteorological variables contained in this data set is listed here:

e0	Potential evaporation for open water surface
es	Potential evaporation for bare soil
et	Potential Evapotranspiration for reference crop
pd	Vapour pressure
pr	precipitation
rg	Calculated radiation
ta	Mean daily temperature
tn	Minimum temperature
tx	Maximum temperature
ws	Wind Speed

### Data period:

The data set contains daily maps for the time period 01/01/1990 (e.g. `et000000.001`) to 31/12/2011 (e.g., `et000008.035`).

### Coordinate system:

The Lambert Azimuthal Equal Area Coordinate Reference System is used as common reference system of this data set. The GISCO Lambert Azimuthal Equal Area projection is characterized by the following parameters:

<i>Projection Name:</i>	GISCO_LAEA
<i>Projection Type:</i>	Lambert Azimuthal Equal Area
<i>Spheroid:</i>	Sphere
<i>Radius of sphere of reference:</i>	6378388
<i>Units:</i>	meters
<i>Longitude of centre of projection:</i>	09° 00' 00"
<i>Latitude of centre of projection:</i>	48° 00' 00"
<i>False easting:</i>	0.0
<i>False northing:</i>	0.0

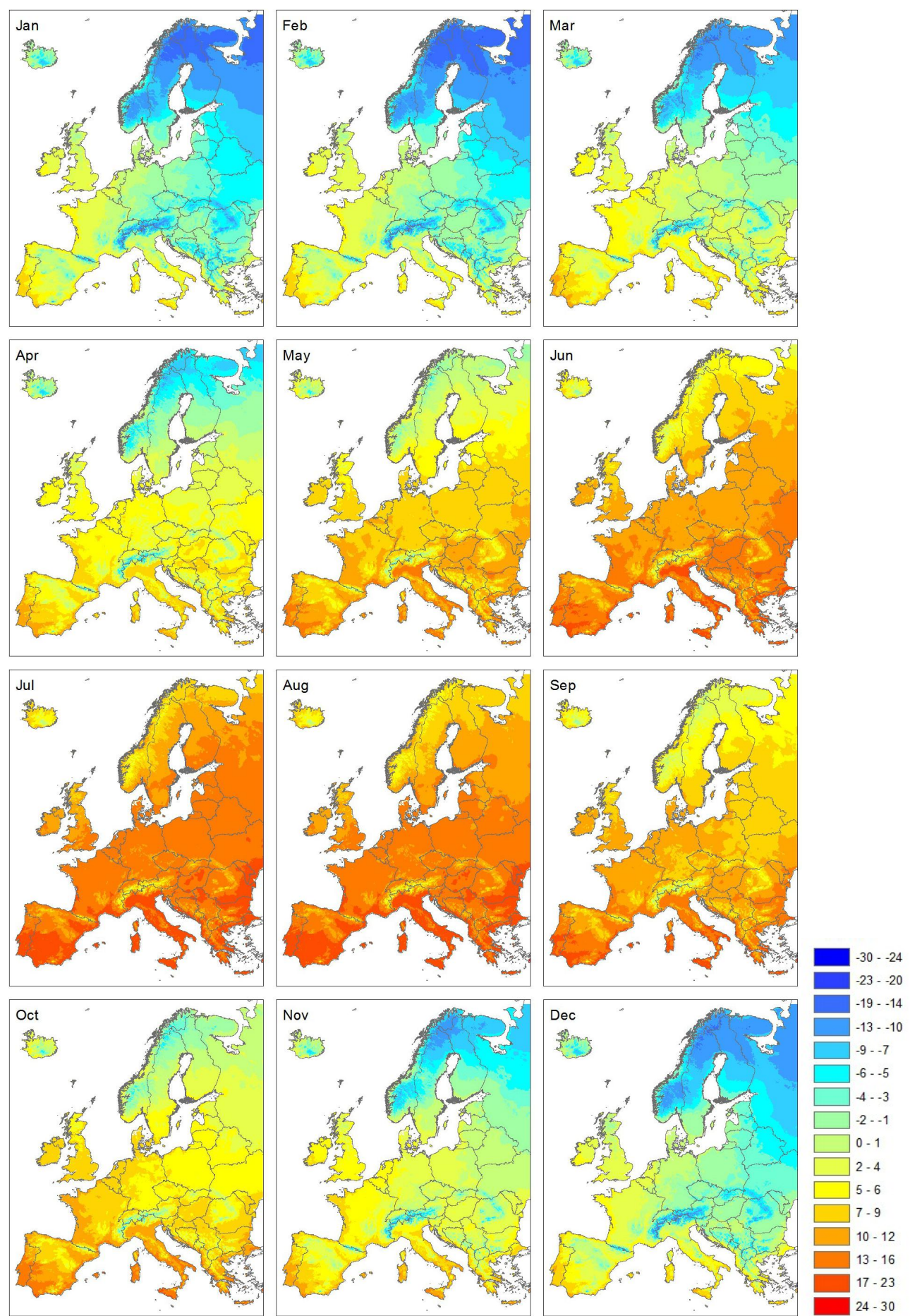
When changing the spatial reference it is recommended to use the command line utility `gdalwarp` from the Geospatial Data Abstraction Library (GDAL - <http://www.gdal.org/>). An example command line transforming a raster map from GISCO\_LAEA to the ETRS\_1989\_LAEA coordinate system is given below:

```
gdalwarp -t_srs "+proj=laea +lat_0=52 +lon_0=10 +x_0=4321000.0 +y_0=3210000.0 +a=6378137.0 +b=6356752.3141403561 +units=m +no_defs " -s_srs "+proj=laea +lat_0=48 +lon_0=9 +x_0=0 +y_0=0 +a=6378388 +b=6378388 +units=m +no_defs " -ot Float32 -tr 5000 5000 <source map name> <output map name>
```

### Data providers:

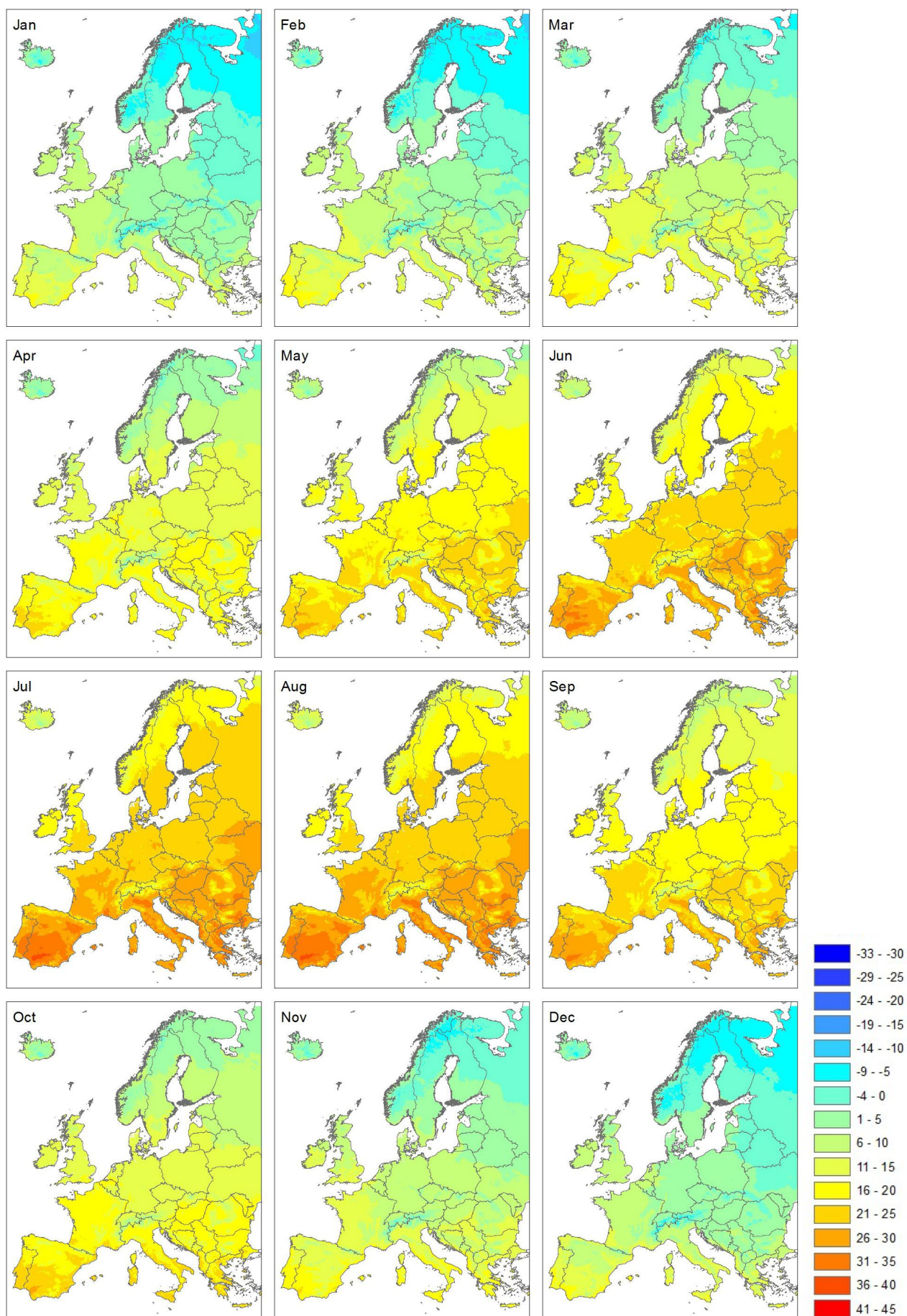
The following table provides an overview of all the different data providers, which have been used to generate the EFAS-Meteo dataset:

Climatology of EFAS-Meteo:



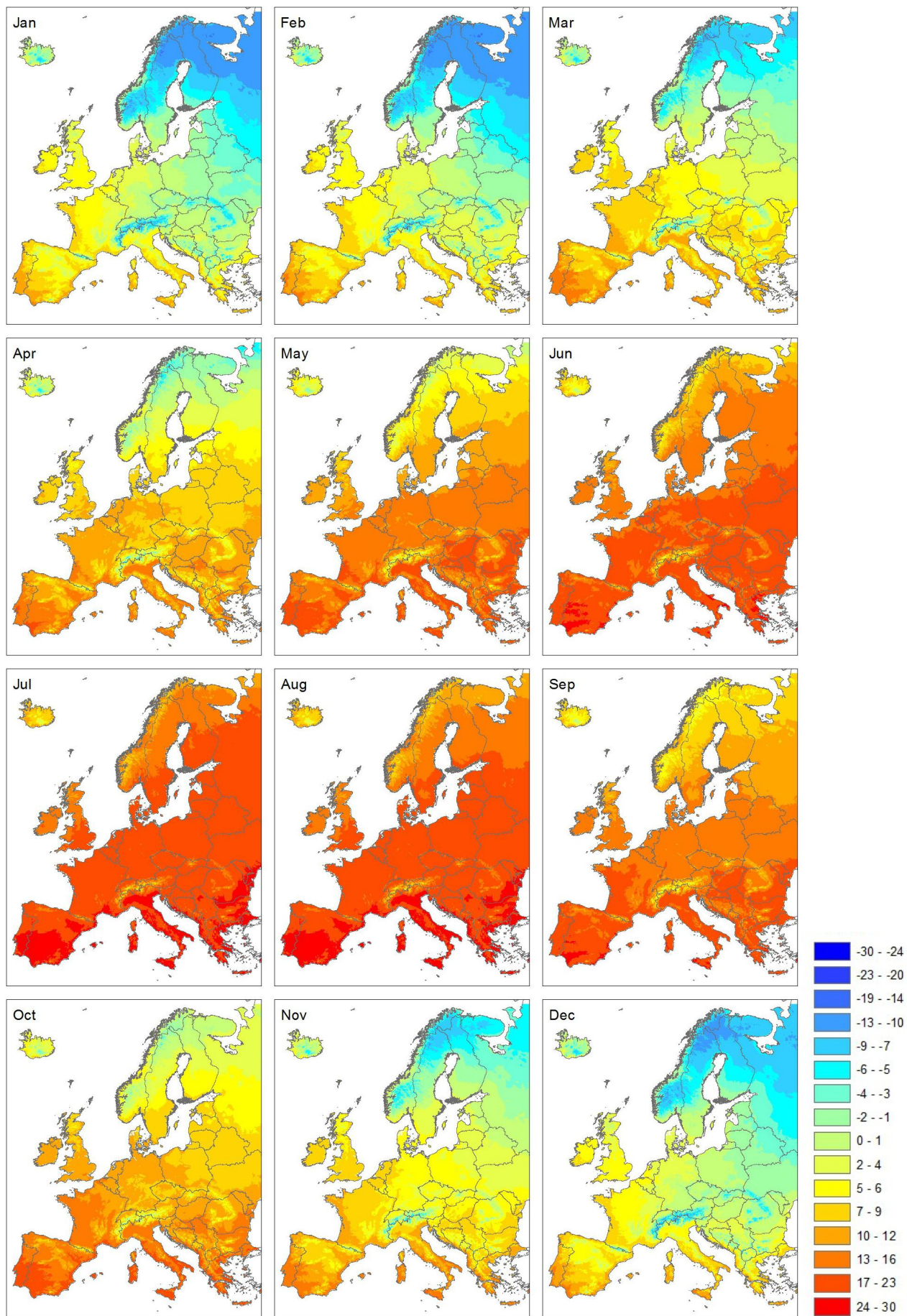
Mean monthly minimum temperature (°C) for period 1990-2011 from EFAS-Meteo





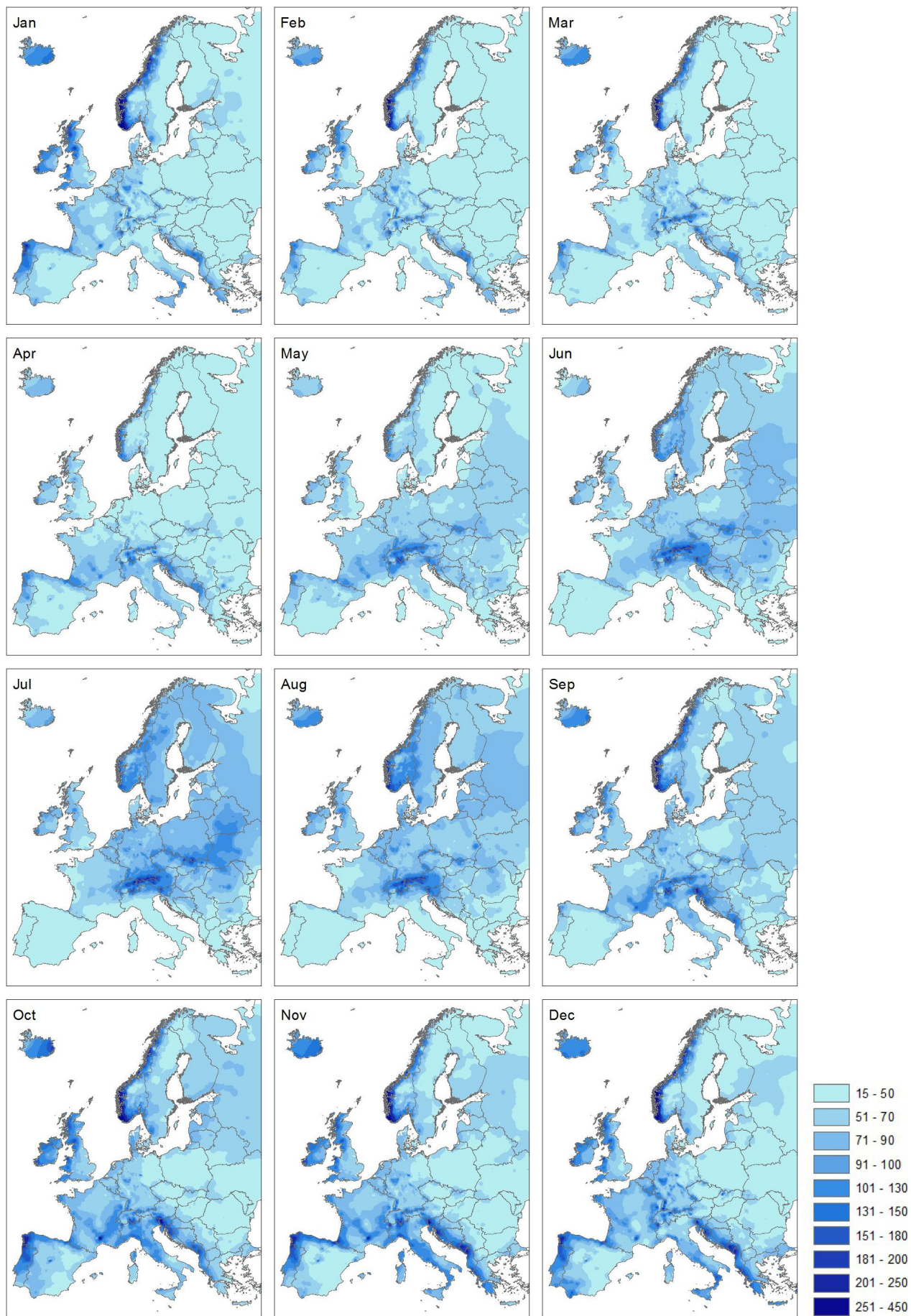
Mean monthly maximum temperature (°C) for period 1990-2011 from EFAS-Meteo



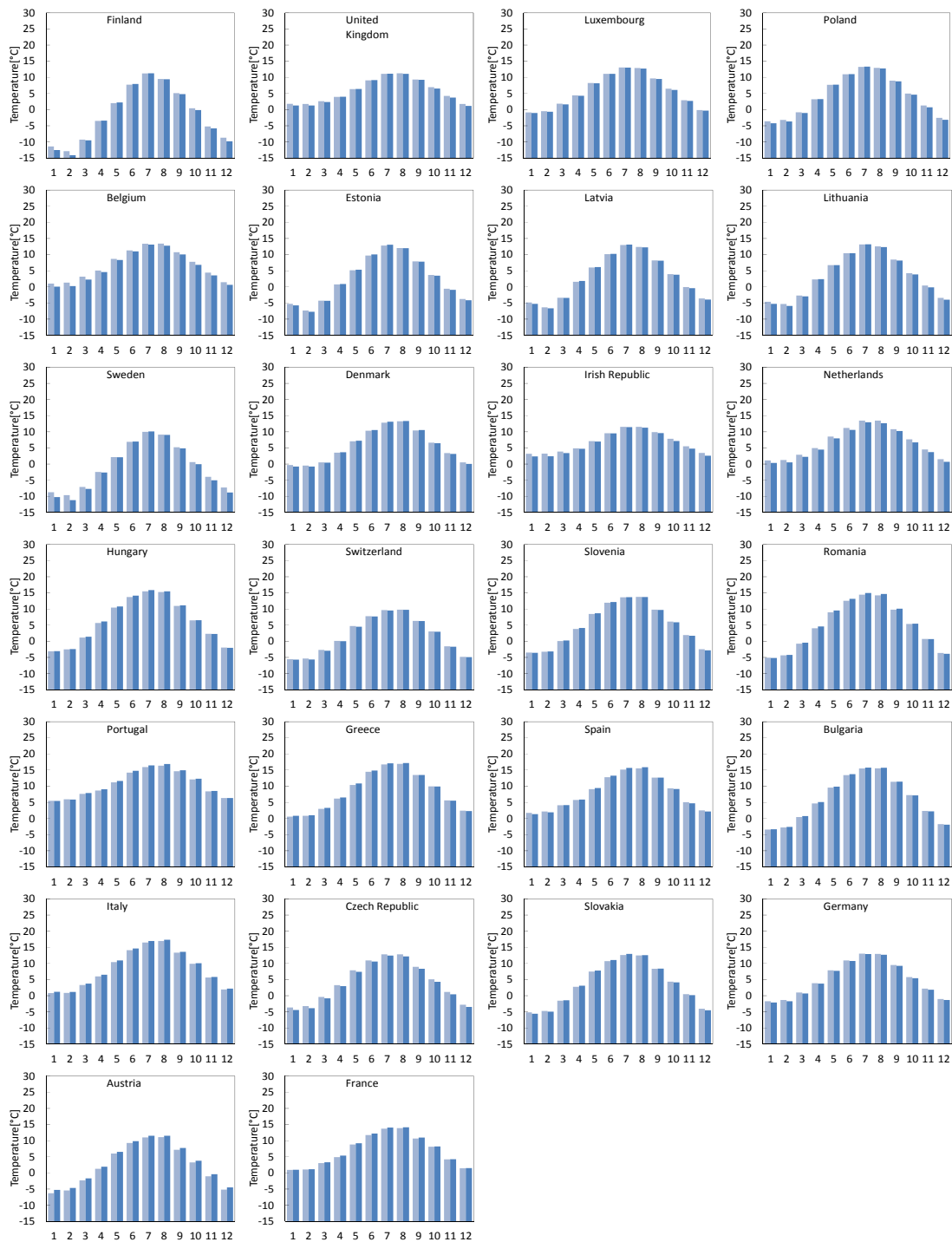


Mean monthly daily mean temperature (°C) for period 1990-2011 from EFAS-Meteo

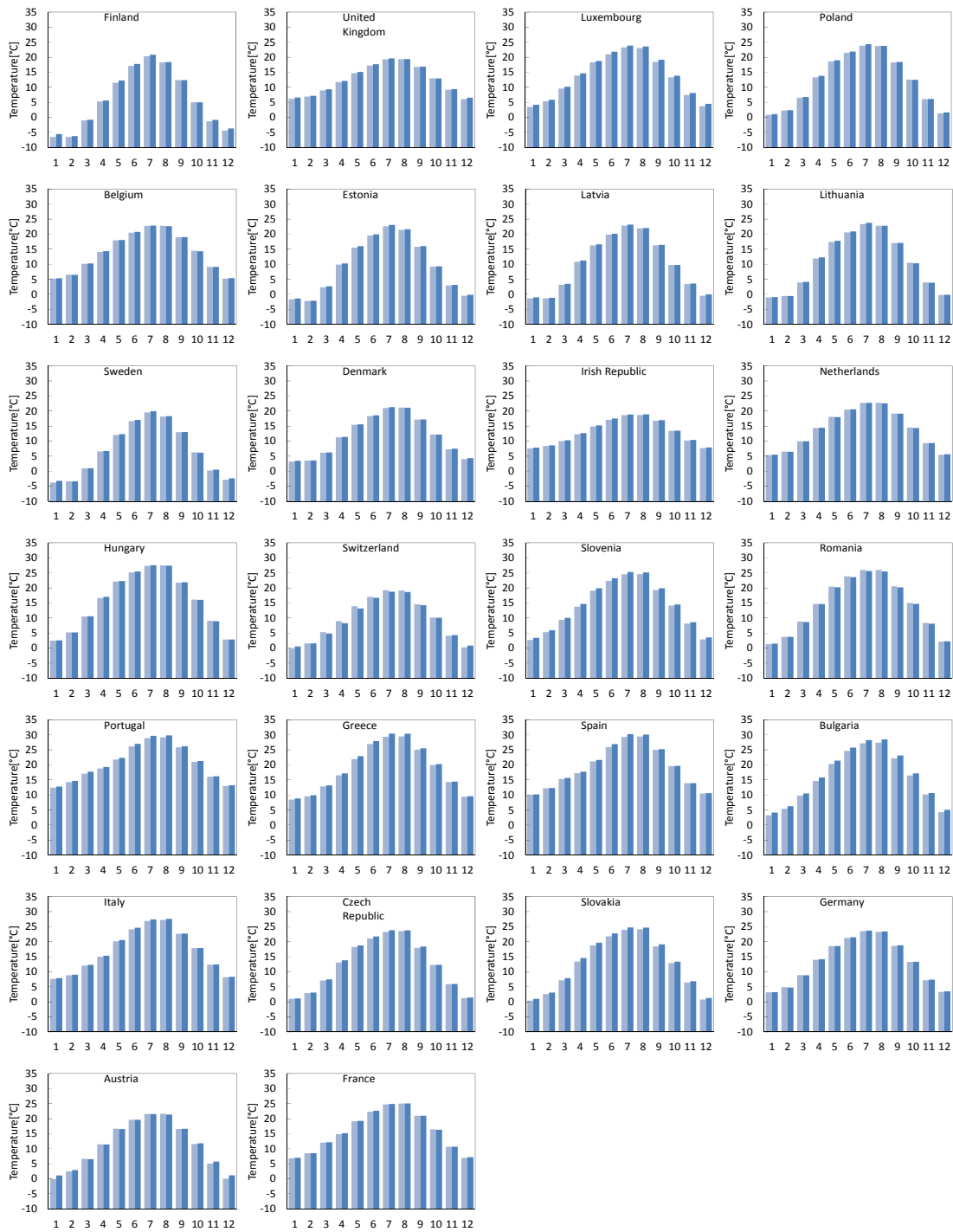




Mean monthly precipitation (mm) for period 1990-2011 from EFAS-Meteo

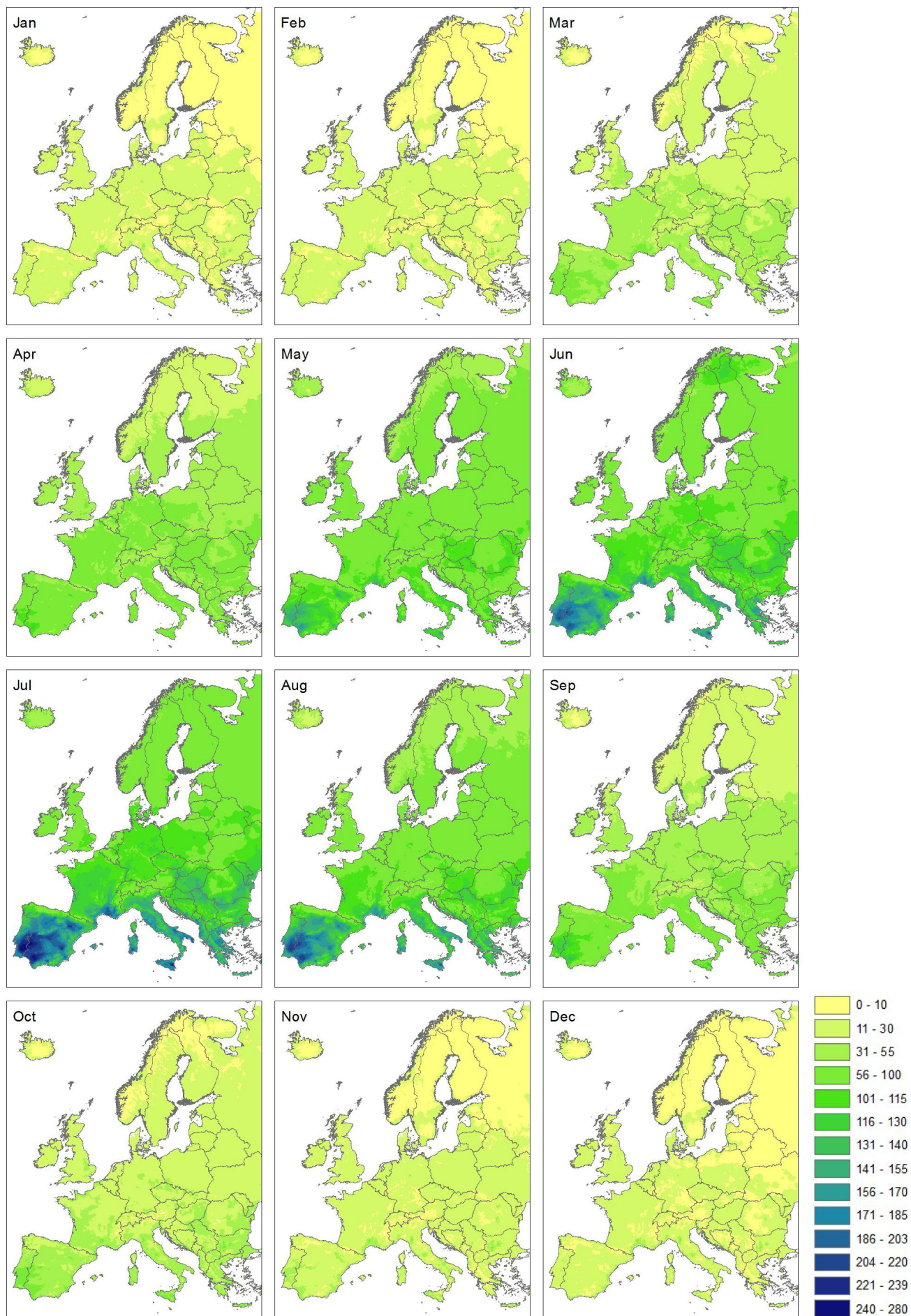


Mean monthly minimum temperature for EFAS-Meteo (light) and E-OBS(dark) for 1990-2011



Mean monthly maximum temperature for EFAS-Meteo (light) and E-OBS(dark) for 1990-2011





Mean monthly potential evapotranspiration (ET in mm) from EFAS-Meteo for the period 1990-2011



European Commission

**EUR 26408 – Joint Research Centre – Institute for Environment and Sustainability**

**Title: EFAS-Meteo: A European daily high-resolution gridded meteorological data set for 1990 – 2011**

**Author(s):** V. Ntegeka, P. Salamon, G. Gomes, H. Sint, V. Lorini, M. Zambrano-Bigiarini, J. Thielen

Luxembourg: Publications Office of the European Union

2013 – 46 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series –ISSN 1831-9424 (online)

ISBN 978-92-79-35006-1 (pdf)

doi: 10.2788/51262

## **Abstract**

Data sets of spatially irregular meteorological observations interpolated to a regular grid are not only important for climate analyses but are also essential in order to derive climatologies for rainfall-runoff models which require meteorological data sets as input forcing. For example, in the European Flood Awareness System long term observed meteorological data are used to drive the hydrological model LISFLOOD to obtain long term time series of simulated discharges at a pan-European scale. Those long term time series of simulated “proxy” discharges can then be used for statistical analysis, e.g., to derive return periods or other time series derivatives.

In this report, we present a comprehensive pan European high-resolution gridded daily data set (EFAS-Meteo) of precipitation, surface temperature (mean, minimum and maximum), wind speed, vapour pressure, calculated radiation and evapotranspiration (potential evapotranspiration, bare soil and open water evapotranspiration). The data set was created as part of the development of EFAS and has been continuously updated throughout the last years.

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

